



STATE: PA  
 COUNTRY: USA  
 ZIP: 19482  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Disquette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEC for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/986,485  
 FILING DATE: 08-DEC-1997  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/059,448  
 FILING DATE: 22-SEP-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: PRESTIA, PAUL F  
 REGISTRATION NUMBER: 23,031  
 REFERENCE/DOCKET NUMBER: GI-70264  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 610-407-0700  
 TELEFAX: 610-407-0701  
 TELEX: 846169  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1101 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-986-485-2

Query Match 10.9%; Score 408; DB 3; Length 1101;  
 Best Local Similarity 21.8%; Pred. No. 1.5e-28; Mismatches 277; Indels 264; Gaps 30;  
 Matches 185; Conservative 124; Mismatches 277; Indels 264; Gaps 30;

QY 8 LLLAVV-----AGATATVWVPPWHVPCPPQACOIRPVWTPRSSYREATIVDCN 56  
 Db 18 LLLIWLILLRLEPVTAAGPRA-----PCAACTCAGDPCTCAGDS-----LDCG 62  
 QY 57 DLFLUTAVPPALPACITQTLIQLQNSNIVRVDQSELGLYLANLTE----- 97  
 Db 63 GRGLALAPLPSDLPSPWTRSLNLSYKHAELDPNQEVYLNHELTAVASIGAGSS 122  
 QY 98 -----LDSLSONSPSDARDCDFHALPOLSLHNEENL 129  
 Db 123 QWVALLFLQQQNRSLDGSGQKAYISLVEVDLNLNNITEVNTYFFHGPPIKELNAGNRI 182  
 QY 130 TRLEPHSFAGLA-SIQELYINHNOYLRIAPRAFGSGLSNLLRHLNSNLLRAIDSFRWEML 188  
 Db 183 GTLGLGAFDGLSRSLLTRLSKRNTRTOLVRAFK-LPRLTQDLDLNRRRLIGETFOGL 241  
 QY 189 PNLTSLIMGGNKVDAILDNFRPLANLRLIVLAGMNLRLRISDVALEGOSLESISFYDQ 248  
 Db 242 NSLEVVKLQRNNISKLTDGAFWGLSKMHVHLLEYDSLEVNSGSLYGLPQALHQHLSNS 301  
 QY 249 LARVFR-----ALEQQPGKFKFLDINKNPKLQRVGPDDFAMN 284  
 Db 302 IARTRKGMWSFCQKILHELVLSFNNTLRLDEESLAELSSSVLRLSHNSHIAEGAKGL 361  
 QY 285 LHLKELGGLANNEELWSID---KFALVNLPELTKLIDTNNPRLSFTHPRFHLPQMDPL 340  
 QY 362 RSLRVLDDHNEISGTIEDTSGAEGLEFGH-SKRLTFLGN-KIKSVKRAFSGLEGLEHL 419  
 QY 341 MNNNHALSALHQVQVESLNLQEWGLHGNPIRCDCVIRW-----ANAT---- 383  
 Db 420 NLGGNAIRSVQFDAFKMKNLKEHISDFLDCQKLWLPMLGRMLQAFVTTACHP 479  
 QY 384 ---GTRVRFIEPOSTLC-----AEP----- 400  
 Db 480 ESLKGQSTFSVPPESFVCDFLRQPIQITQETMMAGKDIRTCSSAASSSSSPMTEAWK 539

RESULT 4  
 US-08-190-8802A-49  
 Sequence 49, Application US/08190802A  
 ; Patent No. 5519003  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mochly-Rosen, Daria  
 ; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
 ; NUMBER OF SEQUENCES: 265  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Dehlinger & Associates  
 ; STREET: P.O. Box 60850  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94306-0850  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/190,802A  
 ; FILING DATE: 01-FEB-1994  
 ; CLASSIFICATION: 530  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Fabian, Gary R.  
 ; REGISTRATION NUMBER: 33,875  
 ; REFERENCE/DOCKET NUMBER: 8600-0139  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 324-0880  
 ; TELEFAX: (415) 324-0960  
 ; INFORMATION FOR SEQ ID NO: 49:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 605 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; INDIVIDUAL ISOLATE: Insulin-like growth factor binding  
 ; INDIVIDUAL ISOLATE: protein complex, Fig. 32  
 ; US-08-190-802A-49

Query Match 10.9%; Score 407.5; DB 1; Length 605;

RESULT 6  
 US-09-905-088a-245.rai  
 ; Sequence 49, Application US/08473089  
 ; Patent No. 6542368  
 ; GENERAL INFORMATION:  
 APPLICANT: Ron, Dorit  
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses  
 TITLE OF INVENTION: Thereof  
 NUMBER OF SEQUENCES: 265  
 CORRESPONDENCE ADDRESS:  
 ADDRESSE: Morrison & Foerster  
 STREET: 2000 Pennsylvania Avenue, NW  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20006-1812  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/473,089  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: MURASIGE, KATE H.  
 REGISTRATION NUMBER: 29,959  
 REFERENCE/DOCKET NUMBER: 2550-0025.22  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 887-1500  
 TELEFAX: (202) 887-0763  
 INFORMATION FOR SEQ ID NO: 49:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 605 amino acids  
 TYPE: amino acid  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 INDIVIDUAL ISOLATE: Insulin-like growth factor binding  
 INDIVIDUAL ISOLATE: protein complex, Fig. 32  
 US-08-473-089-49

Query Match 10.9%; Score 407.5; DB 4; Length 605;  
 Best Local Similarity 25.2%; Pred. No. 6.1e-29;  
 Matches 152; Conservative 72; Mismatches 173; Indels 205; Gaps 16;

QY 3 LLVPLPLLLAWA-----GATANPVVWVWHVCPQQCACQIRPWYTPRSSY--REATT 52  
 Db 8 LALALLLWSVALGERSLGGADPGMPGEAEGPACPAACVC-----SYDDADELS 57

QY 53 VDCMDELTAVPPALPAGTQTLIQLQNSLVRV-----DQSELEGY----- 91  
 Db 58 VFCSSRNRLRLPDCVPGQGQALWIDGNNLSSVPFAFQNLSSLGFLNLOGQQLGSLEFOA 117

QY 92 ---LANTEDLSDONFSNDARCDCFDHALPOLLSHLEENVOLTRI----- 133  
 Db 118 LIGLENLCHLHLERNQLRSALGTFFAHTPALASLGSLNSRLSRLEDGLFEGLGLSLWDLNL 177

QY 134 -----DHSFAGGLASQELVYLNHQYRIPRAFRSLSNLRLHNSNLRAIDSRW 184

Db 178 GWNSLAVIPDAFRGLSRLVLAGNRLAYLQDPLAFLSGLAELRELDLSRNALRAIKANV 237

QY 185 FEMPNLEILMIGGNKVDAI-----LDMN----- 208  
 Db 238 FVQLPRLQKLYDRLRNLLAVALVAPGAFGLKLKALRMWDLSHNRVAGLEDFFPGLGLGVRL 297

QY 209 -----FRPLANLRLSVALGMNLREISDVALEGQLOSLESLSFYDNOLARV--- 252

RESULT 7  
 US-09-063-950-2  
 Sequence 2, Application US/09063950C  
 ; Patent No. 6225085  
 ; GENERAL INFORMATION:  
 APPLICANT: Holtzman, Douglas A.  
 TITLE OF INVENTION: NOVEL LRG PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
 TITLE OF INVENTION: THEREFOR  
 FILE REFERENCE: MEI-019  
 CURRENT APPLICATION NUMBER: US/09/063,950C  
 CURRENT FILING DATE: 1998-04-21  
 NUMBER OF SEQ ID NOS: 9  
 SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO 2  
 LENGTH: 673  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-063-950-2

Query Match 10.8%; Score 402; DB 4; Length 673;  
 Best Local Similarity 24.9%; Pred. No. 2.3e-28;  
 Matches 194; Conservative 73; Mismatches 306; Indels 206; Gaps 24;

QY 4 LYVPLLLAWAGATATVPPVWVWHVCPQQCACQIRPWYTPRSSYREATTVDVCDNLFATV 63  
 Db 7 LLLPLPLLLALG-----PVGQCPGSCOC-----QDQTVCFQARQGY 46

QY 64 PPAIPAGTQTLIQLQNSLVRV-----DQSELEGY----- 123  
 Db 47 PRDPPDTVGLYVYFENGTMIDAGSPAGLPGQLDIDSONQIAASLPSVQFQPLANLSD 106

QY 124 LEENQNLTRLEDDHSFAGLQELVYLNHQYRIPRAFRSLSNLRLHNSNLRAIDSRW 183  
 Db 107 LTANRLHETNETFGRRLRLRVLQGKRNKIRHQGAFDTDLRLLKQDNLRLAPPL 166

QY 184 FWEMLPNLFTIMIGGNKVDAI-----LDMN----- 243  
 Db 167 --RPLRLLDLSHNSLRLAEGPLTANVRLRAGLQQLDLSFLSFLRNHL 222

QY 244 FYDNQLARVPRRALEQVPGIQLFLDLNKNPLQRYVPGDFANMLHKLKGJLNNHPELVSID 303  
 Db 223 VSDQMLERV-----PVIR----- 236

QY 304 FALVNLPPELTKPDITNNPRSFHPRAFHFLPOMETMLNNNALSALHQTYESLPNQE 363  
 Db 237 ---GLRGITRLLAGNTRIAQLRPEDLAGLALQEDVSNISLQLPGLSGLFPLRLL 292

QY 364 VGLHGNP1RCDCVIRWAN--AGTCRVRFIPOSTLCAEPPDILORLPREVPREM---- 416



Db 241 LPRIQKLYLDRNLITAVCPAGAFLGKALRMLDLSHNRVAGLMDFFPGGLGLHVLRLAHN 300 Qy 209 -----FRLANLURSLVLAGMNLRETDYALUGLQSLSESLSYDNLARV----- 252 Db 301 AIAASLRPRFKDHLHLEELQQLGHNRKIROGLERTFEGLQOLEVLNTNDQTEVRVAGFSG 360 Qy 253 -----PRRALEQPGKLFKDLINKNPLQRVGPQDFANMHLKEGLNNM 295 Db 361 LENVAVMNLSGNCLSLPVERFOGLDKLHSLLHSLGHVLHFFAGLGLRFLR- 419 Qy 296 ELYSIDKALVNLFELTKDITNNPRLSIHPRAFHLPOMETMLNNHNSAL----- 350 Db 420 NSISSIEGLSLAGLSELELDITN-RLTHLPROLQFGLCHLEYLLSINQNLTSIAEV 478 Qy 351 -----HO-----QTVESLPLNQEVGLHGNP 372 Db 479 GPLORAFWLDISHNHLLETALAGLFLSSLGRVYRLSLRNNSLTSFSPQGLERLWIDANPND 538 Qy 373 CDCVYR-----WANATGTRVRFEP-----OSTLCAEPPDLQRLPYREVPR 414 Db 539 CSCPLKLALRDFALONGVUVPRFVOTVCESDDCQPVYTYNNITCAGGPNVSGLDLFDVSET 598 Qy 415 EMTDHIC 420

RESULT 13

; Sequence 2, Application US/09191647

; Patent No. 6046015

; GENERAL INFORMATION:

; APPLICANT: Goodman, Corey

; APPLICANT: Kid, Thomas

; APPLICANT: Brose, Katja

; APPLICANT: Tessier-Lavigne, Marc

; TITLE OF INVENTION: Modulating Robo: Ligand Interactions

; FILE REFERENCE: B98-031-3

; CURRENT APPLICATION NUMBER: US/09/191,647

; CURRENT FILING DATE: 1998-11-13

; EARLIER APPLICATION NUMBER: 60/065,544

; EARLIER FILING DATE: 1997-11-14

; EARLIER APPLICATION NUMBER: 60/081,057

; EARLIER FILING DATE: 1998-04-07

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 2

; LENGTH: 1525

; TYPE: PRT

; ORGANISM: human

US-09-191-647-2

RESULT 14

; Sequence 2, Application US/09540245A

; Patent No. 6270984

; GENERAL INFORMATION:

; APPLICANT: Goodman, Corey

; APPLICANT: Kid, Thomas

; APPLICANT: Brose, Katja

; APPLICANT: Tessier-Lavigne, Marc

; TITLE OF INVENTION: Modulating Robo: Ligand Interactions

; FILE REFERENCE: B98-031-3

; CURRENT APPLICATION NUMBER: US/09/540,245A

; CURRENT FILING DATE: 2000-03-31

; PRIORITY NUMBER: 60/065,544

; PRIORITY FILING DATE: 1997-11-14

; PRIORITY APPLICATION NUMBER: 60/081,057

; PRIORITY FILING DATE: 1998-04-07

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 2

; LENGTH: 1525

; TYPE: PRT

; ORGANISM: human

US-09-540-245A-2

Query Match 8.9%; Score 331; DB 3; Length 1525; Best Local Similarity 17.8%; Pred. No. 3.4e-21; Matches 126; Conservative 76; Mismatches 177; Indels 330; Gaps 11; Qy 29 CPPOCACQIRPWTPRSSYREATVDCNDLFUTAVPPALPAGTQTLIQLQNSNIVRVDPOSE 88 Db 28 CPPQSC-----SSSTVDCHGLAKSVPNIPRINTERDLNGNNITRITKD 74 Qy 89 LGYLANLTEDLDLSONSFSDAROCDPHALPOLSLHLEENQLTRLEDSFAGLASLQELYL 148 Db 75 FAGLRLHVLQMLMENKLISTIERGAQDKELERLRLRNHLQFLPELFLGAKLYRLDL 134 Qy 149 NHNOLYRIAPRAESGLSMLRLHNSLRAIDSWEMLPNELIMIGGGNKVDAI---- 204 Db 135 SENQIAQIPKAFQGAVDINKNQDYNQISCTEDGAFLRDLLEVLTNNNNITRLSVAS 194 Qy 205 -----

RESULT 13

; Sequence 2, Application US/09191647

; Patent No. 6046015

; GENERAL INFORMATION:

; APPLICANT: Goodman, Corey

; APPLICANT: Kid, Thomas

; APPLICANT: Brose, Katja

; APPLICANT: Tessier-Lavigne, Marc

; TITLE OF INVENTION: Modulating Robo: Ligand Interactions

; FILE REFERENCE: B98-031-3

; CURRENT APPLICATION NUMBER: US/09/191,647

; CURRENT FILING DATE: 1998-11-13

; EARLIER APPLICATION NUMBER: 60/065,544

; EARLIER FILING DATE: 1997-11-14

; EARLIER APPLICATION NUMBER: 60/081,057

; EARLIER FILING DATE: 1998-04-07

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 2

; LENGTH: 1525

; TYPE: PRT

; ORGANISM: human

US-09-191-647-2

Query Match 8.9%; Score 331; DB 4; Length 1525; Best Local Similarity 17.8%; Pred. No. 3.4e-21; Matches 126; Conservative 76; Mismatches 177; Indels 330; Gaps 11; Qy 29 CPPOCACQIRPWTPRSSYREATVDCNDLFUTAVPPALPAGTQTLIQLQNSNIVRVDPOSE 88 Db 28 CPPQSC-----SSSTVDCHGLAKSVPNIPRINTERDLNGNNITRITKD 74 Qy 89 LGYLANLTEDLDLSONSFSDAROCDPHALPOLSLHLEENQLTRLEDSFAGLASLQELYL 148 Db 75 FAGLRLHVLQMLMENKLISTIERGAQDKELERLRLRNHLQFLPELFLGAKLYRLDL 134









PR	15-MAR-2000; 2000WO-US05884.
PR	30-MAR-2000; 2000WO-US08439.
PR	17-MAY-2000; 2000WO-US13705.
PR	PA (GETH ) GENENTECH INC.
DB	210 nfkplanrlsrlvlagmyltdipgavglgslesfynklykvpqqlqkvpnklfd 269
QY	268 LNNKNPLQRGGDFAANMLKKEGLGNMMEELVSTDKFAVNLNPLTKLIDTNPRLSFIH 327
DB	270 lnkpnkliqegdkmlrlkeglnnmgelvsdryalndnlpelktieatnlpklyh 329
QY	328 PRAFHPLQMETLMNNNALSHQTYESLPLNUQEVGHGPNPIRCDVIRWANATGTRV 387
DB	330 rlafrsypalesimlnnlnalnaiyqktvlesqplmishsfprlnvdigttvfdcr 389
QY	388 REPEPOSTLCAEPPDQLQRLPVREPRETMHCLPLISPRFSPSLOVASESMVHLHRA 447
DB	450 maepepeiywvtpbignktvvetisdkyklissegtleisnqledsgrycvaqnvgadt 509
QY	508 KTVVWVGRALIOPGRDGGQGLRVRQHPTVHILLSWWTPMVSTNLWSSAS-SLRG 566
DB	510 rvatikvngtl---dgtqvlykivqktvlesqplmishsfprlnvdigttvfdcr 565
QY	567 OGATALARLPRGHNSYNTIRLQTEWVACLOVAFADAHOLACVWAKTEA----- 618
DB	566 phitytarvpvdyheynthlqstdjrcvcltvsninqtqkscvnyttkaafavd 625
QY	619 ---TSCHRLGDRGLTIALAV 639
DB	626 getstalaavmgsmfavislasiav 650
RESULT	7
ID	AAB50965 standard; Protein; 716 AA..
AC	AAB50965;
XX	XX
DT	21-MAR-2001 (first entry)
DE	Human PRO1338 protein.
KW	Human; PRO; cytostatic; neuroprotective; respiratory general; antiinflammatory; nootropic; neuroprotective; immunosuppressive; immunostimulant; PROagonist; cancer; inflammatory disorder; immunological disorder; Homo sapiens.
OS	XX
PN	W0200073348-A2.
XX	XX
PD	07-DEC-2000.
XX	30-MAY-2000; 2000WO-US14941.
PF	XX
PR	02-JUN-1999; 99WO-US12252.
PR	22-JUN-1999; 99US-0140650.
PR	23-JUN-1999; 99US-0141037.
PR	20-JUL-1999; 99US-0144758.
PR	01-SEP-1999; 99WO-US20111.
PR	08-SEP-1999; 99WO-US20594.
PR	29-OCT-1999; 99US-0162506.
PR	30-NOV-1999; 99WO-US28313.
PR	01-DEC-1999; 99WO-US28314.
PR	02-DEC-1999; 99WO-US28551.
PR	16-DEC-1999; 99WO-US30055.
PR	20-DEC-1999; 99WO-US30999.
PR	06-JAN-2000; 2000WO-US0376.
PR	11-FEB-2000; 2000WO-US03165.
PR	18-FEB-2000; 2000WO-US04341.
PR	18-FEB-2000; 2000WO-US04342.
PR	02-MAR-2000; 2000WO-US05841.
PR	03-MAR-2000; 2000US-0187202.
PR	10-MAR-2000; 2000WO-US06319.
Query Match	45.4%; Score 1695; DB 22; Length 716;
Best Local Similarity	53.6%; Pred. No. 3.5e-127;
Matches	335; Conservative 104; Mismatches 166; Indels 20; Gaps 5
QY	29 CPPQACACIRPWTPRSSYREATTVCNDLFLAVPAPALPASTQTLQNSIVR VDQS 87
DB	32 cpqicveirpftpqstyreattdvndrlritripsonlsstqvlqllqsniakevd-- 89
QY	88 EGYLANNITELDISQNSFSDARCDCDFHALPOLISLHIZENQLTRLEHDFASGLASOLEY 147
DB	90 elqqlnltedqsfqntnirevgantqtlitneqntemtayclqlsnitqly 149
QY	148 LHNQNLTRIAPRFAFGSINLULHNSNLRAIDSRAFEMUNLETIMIGKNUDILM 207
DB	150 inhnqstisahafagklrlhlnsnklyvidsrtwdfastpnlleimgnpvcild 209
QY	208 NERPLANSLVLAGMURESDYALEEGLQESLESLSDNQARVRALEVPGKFLD 267
DB	210 nfkplanrlsrlvlagmyltdipgavglgslesfynklykvpqqlqkvpnklfd 269
QY	268 LNNKNPLQRGGDFAANMLKKEGLGNMMEELVSTDKFAVNLNPLTKLIDTNPRLSFIH 327
DB	270 lnkpnkliqegdkmlrlkeglnnmgelvsdryalndnlpelktieatnlpklyh 329
QY	328 PRAFHPLQMETLMNNNALSHQTYESLPLNUQEVGHGPNPIRCDVIRWANATGTRV 387
DB	330 rlafrsypalesimlnnlnalnaiyqktvlesqplmishsfprlnvdigttvfdcr 389
QY	388 REPEPOSTLCAEPPDQLQRLPVREPRETMHCLPLISPRFSPSLOVASESMVHLHRA 447
DB	390 rmeplsmfcamppeykgkgkevliqdsseqplmishsfprlnvdigttvfdcr 449
QY	448 LAEPEPEIYWTPTAGCRLTPAHGRYRVRVYEGTELELRVTAEEAGLYTCVAQNLVGDAT 507
DB	450 maepepeiywvtpbignktvvetisdkyklissegtleisnqledsgrycvaqnvgadt 509
QY	508 KTVVWVGRALIOPGRDGGQGLRVRQHPTVHILLSWWTPMVSTNLWSSAS-SLRG 566
DB	510 rvatikvngtl---dgtqvlykivqktvlesqplmishsfprlnvdigttvfdcr 565
QY	567 OGATALARLPRGHNSYNTIRLQTEWVACLOVAFADAHOLACVWAKTEA----- 618







Search completed: August 31, 2002, 14:36:09  
Job time: 71 sec

PRO polypeptides, to link bioactive molecules to cells expressing PRO polypeptides, to modulate biological activities of cells expressing PRO polypeptides, and to detect the presence of mammalian lung, colon, breast, prostate, rectal, cervical or liver tumours by comparing PRO polypeptide expression in a cell sample to that in a control sample. Some of the 275 sequences are also useful to stimulate the release of tumor necrosis factor-alpha (TNF-alpha) from human blood, the proliferation or differentiation of chondrocytes, the proliferation or gene expression in pericardial cells, the release of proteoglycans from cartilage, the proliferation of inner ear utricular supporting cells or of T-lymphocytes, the release of a cytokine from peripheral blood monocytes (PBMNs), or the proliferation of endothelial cells. Some of the PRO polypeptides may modulate glucose or free fatty acid uptake by skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide to factor VIIA. The PRO polypeptides can be used in assays to identify molecules involved in binding interactions. The polynucleotides encoding PRO polypeptides can be used to generate probes, antisense RNA/DNA, transgenic or knock out animals and can be used in gene therapy.



F;363-386/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>  
 F;387-410/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>  
 F;411-434/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>  
 F;435-458/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>  
 F;483-506/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR17>  
 F;507-529/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR19>

Query Match 10.9%; Score 407.5; DB 2; Length 605;  
 Best Local Similarity 25.2%; Pred. No. 4.7e-20; Mismatches 173; Indels 205; Gaps 16;  
 Matches 152; Conservative 72; Mismatches 173; Indels 205; Gaps 16;

Qy 3 LLVAPLLAWA-----GATATVPVPHVPCPPCACQIRPWPYPRSSY--REBATT 52  
 Db 8 IALALLLISWVALGPRSLLEGADPPTGAEAGPACPAACV-----SYDDADELS 57

Qy 53 VDCNDLELTAVPPALPAGTQTLQLSIVRV-----DQSEBLGY----- 91  
 Db 58 VFCSSRNLTTRPDGPGTQALWQHNLSSIPPAFAFRNLSSLAFLNLLQGQSLRPOA 117

Qy 92 ---LANITELDLSQNSQNSFSDARCDCFHALPOLSLHIFENOTRL----- 133  
 Db 118 LLG-LENLICHHLERNLRSLAGTFAHTPAASLGSLNNRSLRDEGLFEGSLGNLWDLN 177

Qy 134 -----DHSGFAGLASQELYLNHNOLYRITAPRAGSGLSNLRLHLSNLRAIDS 184  
 Db 178 GWNSLAVLPDAAFRGLGSLRELVALGNRATYQPAFGLAELRDLSSRALRAIKAN 237

Qy 185 FEMPLPNEILMIGGKVDAI-----LDNN----- 208  
 Db 238 FVQLPRLQKLYLDRLNIAAVAPGAFGLKALRWDLISHNRYVAGLEDTFPGLGLRVLRL 297

Qy 209 -----FPLANLRSILVLAGMLNLREISDVALEGLOSSESFYDNLARY----- 252  
 Db 298 SHNATAIASLRPRTPKDHFLELQLGHRNRIOLAERSFEGGLQLEVITLDHNOLOEKAGA 357

Qy 253 -----PRRALEQVPGKFLDINKNPLQRVSPGDFANMLHKEGL 292  
 Db 358 FLGLTNWAVMNLSGNCLNLPEQVFRGLGKHLHSLHLEGSLCGRPHFTGSLGRLRFL 417

Qy 293 NNMEELVISDKFALNLPETKLDTINPRPLSFHPRAFHFLPOMETLMNNNAISAL-- 350  
 Db 418 KD-NGLVIGIEBOSLWGLAELLELDLTSN-QLTHLPHRLFQGLGKIEYLLSRRNLAELPA 475

Qy 351 -----HQQTVELPN----- 368  
 Db 476 DALGPIQRAFWLDVSNRHLRPLNPSLALPLGLRLSLRNNSLRFTPQPGLERLWLEG 535

Qy 369 NPIRCDCVIR---WANATGTRV-RFIEP-----QSTLCAEPPDQLPVR 409  
 Db 536 NPWDGCPPLKALRDFALQNSAVPFWQACEGDDCQPPVYVNNITCASPPEVAGLDL 595

Qy 410 EV 411  
 Db 596 DL 597

RESULT 4  
 JC5239  
 1insulin-like growth factor acid-labile chain - baboon  
 C;Species: Papio sp. (baboon)  
 C;Date: 17-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 09-May-1997  
 C;Accession: JC5239  
 R;Deihany, P.; Baxter, R.C.  
 Biochem. Biophys. Res. Commun. 227, 897-902, 1996  
 A;Title: The cloning and expression of the baboon acid-labile subunit of the insulin-like  
 A;Reference number: JC5239; MUID:97040714  
 A;Contents: Liver  
 A;Accession: JC5239  
 A;Molecule type: mRNA  
 A;Residues: 1-605 <DEL>  
 C;Comment: This factor is structurally related to proinsulin and have insulin-like meta

Query Match 10.7%; Score 397.5; DB 2; Length 605;  
 Best Local Similarity 24.9%; Pred. No. 2.3e-19; Mismatches 150; Conservative 73; Mismatches 173; Indels 207; Gaps 16;  
 Matches 150; Conservative 73; Mismatches 173; Indels 207; Gaps 16;

Qy 3 LLVAPLLAWA-----GATATVPVPHVPCPPCACQIRPWPYPRSSY--REBATT 52  
 Db 8 IALALLLISWVALGPRSLLEGADPPTGAEAGPACPAACV-----SYDDADELS 57

Qy 53 VDCNDLELTAVPPALPAGTQTLQLSIVRV-----DQSEBLGY----- 91  
 Db 58 VFCSSRNLTTRPDGPGTQALWQHNLSSIPPAFAFRNLSSLAFLNLLQGQSLRPOA 117

Qy 88 EUGYLANEELTELLSQNSQNSFSDARCDCFHALPOLSLHIFENOTRL----- 133  
 Db 118 LLG-LENLICHHLERNLRSLAGTFAHTPAASLGSLNNRSLRDEGLFEGSLGNLWDLN 176

Qy 134 -----DHSGFAGLASQELYLNHNOLYRITAPRAGSGLSNLRLHLSNLRAIDS 183  
 Db 177 GWNSLAVLPDAAFRGLGSLRELVALGNRATYQPAFGLAELRDLSSRALRAIKAN 236

Qy 184 WREMLPNEILMIGGKVDAI-----LDNN----- 208  
 Db 237 VPAOLPRLQKLYLDRLNIAAVAPGAFGLKALRWDLISHNRYVAGLEDTFPGLGLRVLRL 296

Qy 209 -----FPLANLRSILVLAGMLNLREISDVALEGLOSSESFYDNLARY----- 252  
 Db 297 LSHNATAIASLRPRTPKDHFLELQLGHRNRIOLAERSFEGGLQLEVITLDHNOLOEKAGA 357

Qy 253 -----PRRALEQVPGKFLDINKNPLQRVSPGDFANMLHKEGL 292  
 Db 357 AFLGLTNWAVMNLSGNCLNLPEQVFRGLGKHLHSLHLEGSLCGRPHFTGSLGRLRFL 417

Qy 292 LNMEELVISDKFALNLPETKLDTINPRPLSFHPRAFHFLPOMETLMNNNAISAL-- 350  
 Db 417 KKD-NGLVIGIEBOSLWGLAELLELDLTSN-QLTHLPHRLFQGLGKIEYLLSRRNLAELPA 474

Qy 351 -----HQQTVELPN----- 368  
 Db 475 DALGPIQRAFWLDVSNRHLRPLNPSLALPLGLRLSLRNNSLRFTPQPGLERLWLEG 535

Qy 368 NPIRCDCVIR---WANATGTRV-RFIEP-----QSTLCAEPPDQLPVR 408  
 Db 535 NPWDGCPPLKALRDFALQNSAVPFWQACEGDDCQPPVYVNNITCASPPEVAGLDL 594

Qy 409 REV 411  
 Db 595 RDL 597

RESULT 5  
 T28714  
 Hypothetical protein T21D12.9a - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Sequence\_revision 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
 R;Woessner, J.  
 Submitted to the EMBL Data Library, August 1997  
 A;Reference number: Z2054  
 A;Accession: T28714  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-789 <WQE>  
 A;Cross-references: EMBL:AF016687; PIDN: AAC8096\_1; GSDB:GN00022; CFSN:T21D12.9a  
 A;Experimental source: strain Bristol N2; clone T21D12  
 C;Genes: CFSN:T21D12.9a  
 A;Map position: 4  
 A;Introns: 38/2; 84/2; 132/2; 204/2; 275/2; 351/2; 519/3; 615/1; 714/1; 758/2

Db	178	-----PVQA-----FRSLQALQMTLANKKTHIADAFGNLSSLV 213	Db	465	: :: :: :   --PEPEIWWTPAGLRLT-PAHAGRRYRVYPCTELR 519
Qy	241	SLSPYDNLQALARVPRRALEQVQGLKFLDINKPLQRVGPGDFANMLHKEGLGNMELVS 300	Qy	451	PEPEIWWTPAGLRLT-PAHAGRRYRVYPCTELR 485
Db	214	VLHLNRRHLSLKGKCFDGLHSLETDLNNYNNLDEF-PTAITALSLNKELGFSH-NNIRS 271	Db	520	--EASVVECCSILSKTPER-----IPSTIELR 545
Qy	301	IDKFLALVNLPELTKLDITNNRRLSFTHRPHFLPOMETLMNNNA----- 346			
Db	272	IPERAFTVGNPSLITIHYDNP-IOFGVSAFOHLPRLTLTNGASHITEPHLTGTATL 330			
Qy	347	-----:ISALHOOQVESLPLQEVGLHON 369			
Db	331	ESLTIGAKISSLQAVCDOLPNLQVLDLSYN 362			
RESULT	8		RESULT	9	
T42218		slit-1 protein homolog - rat	Qy		orphan G protein-coupled receptor precursor - human
		N;Alternate names: MBGP4 protein	C;Species: Homo sapiens (man)		C;Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 21-Jul-2000
C;Specie		C;Specie	C;Species: Homo sapiens (man)		C;Accession: JE0176
C;Date:	03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000	R;McDonald, T.; Wang, R.; Bailey, W.; Xie, G.; Chen, F.; Caskey, C.T.; Liu, Q.			
C;Accession:	T42218	Biochem. Biophys. Res. Commun. 247, 266-270, 1998			
R;Nakayama, M.; Nakajima, D.; Nagae, T.; Nomura, N.; Seki, N.; Ohara, O.					
Genomics 51, 27-34, 1998					
A;Title: Identification of high-molecular-weight proteins with multiple EGR-like motifs					
A;Reference number: Z14126; MVID:93360089					
A;Accession: T42218					
A;Status: preliminary; translated from GB/EMBL/DDBJ					
A;Molecule type: mRNA					
A;Residues: 1-1531 <NAK>					
A;Cross-references: EMBL:AB011530; NID:9344989; PID:RAA32460.1; PID:93449290					
A;Experimental source: strain Sprague-Dawley; brain					
C;Genetics:					
A;Gene: MEGF4					
Query Match	9.7%	Score 360; DB 2; Length 1531;	Query Match	9.6%	Score 360; DB 2; Length 907;
Best Local Similarity	23.6%	Pred: No. 3e-16;	Best Local Similarity	27.4%	Pred: No. 1.5e-16;
Matches	136;	Conservative	Matches	134;	Conservative
		Mismatches			Mismatches
		216;			189;
Indels	149;	Gaps	Indels	96;	Gaps
		17;			14;
Qy	8	LLL--AWVAGATAVPVWVPHVPCPPQACQIRPWPYTPRSSYREATVCDNLTLAVP 64	Qy	3	LIVAPILLAWAGATATVPPVPHVPCPPQACQIRPWPYTPRSSYREATVCDNLFTA 62
Db	20	LLMWAQAWRGATA-----CPALCIC-----TGTIVDCHGTLQAI P 56	Db	9	LISLPVLLQIATGGSSPRSGVLLR-GCPTKHC-----PDORMLLRVDCSDGLSE 59
Qy	65	PALPGTQPLLQSQNSIVRVDQSEIGYANLTELDSLONSFSARDCDFHALPQLSLHL 124	Qy	63	VPPALPAGTQTLLOSNSTVYDSEIGYANLTELDSLONSFSARDCDFHALPQLSL 122
Db	57	KNIPTRERLEELNGNINTRIHKNDPAGKQLRVLQVLMENOQIAGAVERGAFFDMKSLERURL 116	Db	60	LPSNLSVFTSYLSDLSNNNISQQLPNPLPSLRFLEELRLAGNALTYIPKGATFTGLYSLKV 119
Qy	125	ERNOITRLDEHSPAGLASQELYLNHNOIYRATPRAFSGLSNLRLHNSNLLRAIDS 184	Qy	123	HLEENOLTRIEDHSPAGLASQELYLNHNOIYRATPRAFSGLSNLRLHNSNLLRAIDS 182
Db	117	NRNQDQVLPQVLFLLFQNQNLSDLSNLSLQVPRKAFRGATDLKNLQDKNQIISCGIEGA 176	Db	120	MQNINOLRHYTEALONLRSLSQSLDNLNTSYVPPSCSGHSLSRHLWMDDNALTEI- 177
Qy	185	FEMFLNLTLMIGGNKVDAITLDMRFRPLANLRLS----- 218	Qy	183	RWFEMLPNLELIMIGGNKVDAITLDMRFRPLANLRLSILLAGMLRBTSDYALLEGIQLES 242
Db	177	FRALIGLEVLITNNNNITIPVSSFNHMPKRLTRLHSNHLFCDCHLNLQSLRQPTI 236	Db	178	-----:PVQA-----FRSLQALQMTLANKKTHIADAFGNLSSLV 215
Qy	219	-----VLAGMLREI-SYALEGIQLESLSFYDNLQALARVPR----- 254	Qy	243	SFYDNLARVPRRALEQVQGLKFLDINKPLQRVGPGDFANMLHKEGLGNMELVS 302
Db	237	GLFTQCSGPASLRGLNVAEVQKSEFSCG-----QGEAAQVPACTLSSCSCPAMC 286	Db	216	HLHNNRHSLSLKGKCFDGLHSLETDLNNYNNLDEF-PTAITALSLNKELGFSH-NNIRS 273
Qy	255	-----RALEQVPG-----LKFLDINKPLQRVGPGDFANMLHKEGLGNMELVS 298	Qy	303	KFLVNLPELTKLDITNNRRLSFTHRPHFLPOMETLMNNNA----- 346
Db	287	SCSGNGIVDQCRGKGGLTAIPANLPETMEIRLEELNGKISIPPAGFSPYKLRRIDLN-NOI 345	Db	274	EKAFTVGNPSLITIHYDNP-IOFGVSAFOHLPRLTLTNGASHITEPHLTGTATL 332
Qy	299	VSTIDKFALVNLPELTKLDITNNRRLSFTHRPHFLPOMETLMNNNA----- 335	Qy	347	-----:ISALHOOQVESLPLQEVGLHON 369
Db	346	AETAPDAFOGLRSLSLNVLYGNKTFIDLPRGVGGLYTQLLNLANKINCIRPDAFOIQ 405	Db	333	LITLTCQAISSLQPTVCNCQIPNLQVLDLSYN-----LLEDLPSFVC 373
Qy	336	QMETLMLNNNALSALHQQTVESLPLQEVGLHGSNPIRCDCVIRWANAGTRVTRIEPOST 395	Qy	398	AEPDLQRLPVPREVPREM-TDHLPLISPRSFPSLQVAGSMEMLHCRALAEPEPIY 456
Db	406	NLSLSLSDNKLQDKSFTSIRAIQIHLAQNPFICDCNLKWL-ADFLRTPNIEITGA 464	Db	374	OK---LOKIDRHNIEYEIKVDPFOQLSLR---SLMLWNKIAITH-----PNAF 418
Qy	396	LCAEPPDQLQLPVPREVPREM-----TDHLPLISPRSFPSLQVAGSMEMLHCRALAE 450			

R;Bodisclair, Y.R.; Seto, D.; Hsieh, S.; Hurst, K.R.; Ooi, G.T.  
 Proc. Natl. Acad. Sci. U.S.A. 93, 10028-10033, 1996  
 A;Title: Organization and chromosomal localization of the gene encoding the mouse acid 1  
 A;Reference number: JG6128; MUID:96413591  
 A;Accession: JG6128  
 A;Molecule type: DNA  
 A;Residues: 1-1603 <B01>  
 A;Cross-references: GB:U66900; NID:91621612; PID:AAH17270.1; PID:g1621613  
 C;Comment: This protein is a serum protein and it is of the ternary complex in the physi  
 C;Genetics:  
 A;Gene: als  
 A;Map position: 17  
 Query Match 9.2%; Score 342.5; DB 2; Length 603;  
 Best Local Similarity 22.0%; Pred. No. 1.3e-15;  
 Matches 139; Conservative 69; Mismatches 190; Indels 233; Gaps 15;  
 QY 8 LILAAWA-----GATATPVPPWVPCPPCACQCPWVTPRSPSYREATVDCN 56  
 Db 13 LLAFFVALGALGFCYQIYDTPGASADE--GPQCPTCTCSY-----DDYTDDELVSFC 61  
 QY 57 DLFLLAVPVLAPLAGTQTLIQLQSNTSVRV-----DSELGY 91  
 Db 62 SRNLQQLPDPGIVPVSRAWLWDGNNLSSIPSAAQNLSSLDFLNQGSWLSPEQALLG- 120  
 QY 92 LANUTELDLSQNSFSDARDCDFHALPOLLSHLENLQNLTRE----- 133  
 Db 121 LQNLVHLHLRNLNLSSLAAGLFRHPPSLASLSLGNLNLGRLEGIFRGLSHLWDNLGNW 180  
 QY 134 -----DHISAGLASLQELTINHNLQYRTPRAFSQNLRLHNSNLRAIDSRRWFM 187  
 Db 181 SLWVLPPTVQGLGLNHLVHLVLAGNKLTYLQPALCGLGELRELIDSRLNSRAVNFH 240  
 QY 188 LPNLLEILMIGGNKVDAI-----LDMN----- 208  
 Db 241 LPRLQKLYLDRNLTVAVAPRAFLGLKRALWDLNLHNRVAGLLEDFTPPGLLGLHVRLAHN 300  
 QY 209 -----FRPLANLRSVLAGMNLRESDYALEGLQSLSESLSYDQNQARV----- 252  
 Db 301 AITSHSPRTEKDLHLEELQOLGHNRNIRQLEKETFGGLQLEVLTDNQIHEVKVGAFFG 360  
 QY 253 -----PRRALEQPGKFLDLNKNPRLQYRGDPFANMLHLKGLGNM 295  
 Db 361 LFNWAVMNLSGNCLESLSPHEVFOGLGRHLHLEHSGLGRHLTAFGLRLFLRD- 419  
 QY 296 EELVSIIDKPALVNLFELTKIDITNNPRLSFTHPRAFHPRHPOMETLMNNNALSALHQ--- 352  
 Db 420 NSISIIEQSLAGLSELLELDLTAN-QLTHLPROFQGIGOLEVLLSNQNLTMSEDVL 478  
 QY 353 -----OTVESLSPNQEVGLHGNPIR 372  
 Db 479 GPLORAFWFLDLSHNRLTEPAEGLFESSLGRYRLNURNNSLQTFVQPLGLRLWILDANPMD 538  
 QY 373 CDCYVIR-----WANATGTRYTRFIR-----OSTLCCEPDPDQQLRPPFREVPR 414  
 Db 539 CSCFLKLARDFALQNPVGPVPRFVOTCEGDCQPVYTYNNITCAGPANYSGLDRDI--- 595  
 QY 415 EMTDHCPLPLSPSFPPLSQAQSESMVHL 445  
 Db 596 -----SETLFVHC 603  
 RESULT 13  
 B6665 slit protein 2 precursor - fruit fly (Drosophila melanogaster)  
 C;Species: Drosophila melanogaster  
 C;Date: 30-Apr-1991 #sequence\_revision 30-Apr-1991 #text\_change 19-May-2000  
 C;Accession: B6665  
 R;Rothberg, J. M.; Jacobs, J. R.; Goodman, C. S.; Artavanis-Tsakonas, S.  
 Genes Dev. 4, 2169-2187, 1990  
 A;Title: slit: an extracellular protein necessary for development of midline glia and cc  
 A;Reference number: A36665; MUID:91099665

A;Accession: B36665  
 A;Title: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-1469 <ROT>  
 A;Cross-references: GB:X63959  
 C;Genetics:  
 A;Gene: FlyBase:sl1  
 A;Cross-references: FlyBase:FBgn0003425  
 C;Superfamily: unassigned EGR-related proteins; EGF homology; leucine-rich alpha-2-glycoprotein repeat homology <PAH1>  
 A;Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
 F:173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
 F:191-220/Domain: proteoglycan carboxyl-terminal homology <PCSL1>  
 F:288-313/Domain: proteoglycan amino-terminal homology <PAH2>  
 F:347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
 F:371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
 F:395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
 F:419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>  
 F:450-494/Domain: proteoglycan carboxyl-terminal homology <PCSL2>  
 F:512-537/Domain: proteoglycan amino-terminal homology <PAH3>  
 F:547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>  
 F:572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>  
 F:596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>  
 F:620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>  
 F:651-695/Domain: proteoglycan carboxyl-terminal homology <PCSL3>  
 F:743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>  
 F:767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>  
 F:846-890/Domain: proteoglycan carboxyl-terminal homology <PCSL4>  
 F:1028-1061/Domain: EGF homology <EGF>  
 F:1068-1099/Domain: EGF homology <EGF2>  
 F:1115-1148/Domain: EGF homology <EGF1>  
 Query Match 8.7%; Score 324; DB 2; Length 1469;  
 Best Local Similarity 24.7%; Pred. No. 9e-14;  
 Matches 103; Conservative 67; Mismatches 169; Indels 78; Gaps 11;  
 QY 29 CPPQACACQIRFWVTPRSSYREATTVDNDLFLTAVPPALPAGTQTLLQNSIVRVDQSE 88  
 Db 519 CPMQCHC-----EGTVDTCGRRKKEPDRDPLHTTELLNDNEGRISCG 565  
 QY 89 IGLYLANTELDLSQNSFSDARDCDFHALPOLLSHLEENQNLTRLDHSFAGLASLQELYL 148  
 Db 566 L-----FGRLPHVKLRLKRNQLTGIEPNATEGASHIQELQ 602  
 QY 149 NHNOLYRIAPRAGSGLSNLRLHNSLRAIDSRRWFEMPLPNLITLMMGNNKVADLM 208  
 Db 603 GENKIKEISNKMFLGLHQLKLTNLNDQIISCCVMPGSFHEHLSLTSINTLASNPNFC---NCH 660  
 QY 209 FRPLAN-LRSVLVAGMLR---EISDYLEGLQSL----- 240  
 Db 661 LAWFAECVRKKSLSANGAARCAGPSKVRDQVQKIDLPHSEKCSSENSEGLGDCGPPSCT 720  
 QY 241 ---SLSFYDNLNQARVPRRALEQPG-LKFLDLNNPKLQRGVPGCDFANMLHLKRLGNM 295  
 Db 721 CTGTVVACSRNQKEIP---GIPAETSELYLESLNEIETQHRYERIRHLRLSLRDLSDN- 775  
 QY 296 EELVSIIDKPALVNLFELTKLDTNNPRLSFTHPRAFHPRHPOMETLMNNNALSALHQTV 355  
 Db 776 NQITLISNYTFANLTLKLSTLISYN-KLQCLQRHALSGLNLLNRLVVSLSHGNRISMLPEGSF 834  
 QY 356 ESLPNLQEVGLHGNPIRCDCVIRWANATGTRYTRFIR-----POSTLCAEPDQQLRPPFREV 412  
 Db 835 EDLKSLSLTHALGSNPLYCDGKWF-SDWIKLQYVEPIARCACPEQOMKDKLILSTP 890

RESULT 14  
 A36665

C;Comment: The amino end of the intact protein is blocked.  
C;Comment: This protein is absent in Bernard-Soulier syndrome.

C;Genetics:

A;Gene: GDB:GP5

A;Cross-references: GDB:230236; OMIM:173511

A;Map position: 5pter-5qter

C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology

C;Keywords: blocked amino end; glycoprotein; platelet; tandem repeat; transmembrane prot

Query Match 8.6%; Score 321; DB 2; Length 560;  
Best Local Similarity 26.6%; Pred. No. 3.5e-14; Mismatches 227; Indels 106; Gaps 16;  
Matches 141; Conservative 57; Mismatches 227; Indels 106; Gaps 16;

Qy 9 LLAWWAGATATPVVFWHVPCPPQACQIRPWYTPRSSYRATVDCNDLF-LTAVPPAL 67  
Db 6 LLCAVAGLGLRQP----FPCPACKC-----VFRDAQCGGGDVARISHL-GL 48

Qy 68 PAG-TQTLIQSNTVRYDQSELQVANLFLDLSQNSFSDARDCCDFHALPOLLHLHE 126  
Db 49 PTNLTHILLFGMGRGVQYQSGSPSG-MTVLQLMISDHISAVAPGTSFDLTKLKLTLR 107

Qy 127 NQLTREIDHSEFAGLQLASQELXLNHMLYRATPRAASGLSNLRLHNSNLRAISRWFE 186  
Db 108 NRITHLPGALLDKMYLLEQOLFDDHNAFLDQKLVNQELALNQNQDFLASFET 167

Qy 187 MLPNNEILMIGGNKVDAILDNFRPIANLRLSVLQAGMLNREISDVALEGIQSLESISFYD 246  
Db 168 NLENFKLDSLGSNNLTHLPKGGLGQAQLERLLISNLRLSLSLNSLIGALTELQFHR 227

Qy 247 NOLARVPRRALEQVQGKFLDNLK-----NPQYVPGDFA 282  
Db 228 NHIRSIAPGAFDRLPNLNLSSLTLSRHLAFLPSALFLSHNLTLTLPENPLAELPGVLFQ 287

Qy 283 NMLHKEGLGNINMELSLVSLKFLNLPETKLDTINNPULSFIIPRAFHLPQMTLML 342  
Db 288 EMGGQELWLNRTQTLTPAERFLRSRERYLGVTLSPRLSALPOGAFFGIGELQVLA 346

Qy 343 NNNASAL-----HQQ-----T 354  
Db 347 HSNGTALPDGLLRLGIGKLRQSLRNRRLRALPRAFLRNLISSLESVQLDNQELPLGDV 406

Qy 355 VESLNLQEGVGLHGPICDC---VIRWANATGTRVREPOSTLCAEPPDLOLULPVE 410  
Db 407 FGALRLTEVLGHNSWRCCGGLGPFLGWLROHLGLGVGEEPR-CAGPGAHAGLPLWA 464

Qy 411 VPERMTDHLQPLPISPRSFPSLQASGESMVLHORALAPEPEPY-WVWP 460  
Db 465 LPGGDA--EC--PGPRGPPRPAADSSEAPVR-PALAPNSSEPWVWAQP 509

Search completed: August 31, 2002, 14:36:40  
Job time: 102 sec



DR	DR	DR	DR
KW	SMART; SM0013; LRR-TYP; 1.	SMART; SM0036; LRR-TYP; 11.	SMART; SM0036; LRR-TYP; 11.
	Glycoprotein; Leucine-rich repeat; repeat; Signal.	Glycoprotein; Leucine-rich repeat; repeat; Signal.	Glycoprotein; Leucine-rich repeat; repeat; Signal.
FT	SIGNAL	1	1
FT	CHAIN	28	27
FT		605	605
REPEAT	53	73	INSULIN-LIKE GROWTH FACTOR BINDING
REPEAT	74	96	PROTEIN COMPLEX ACID LABILE CHAIN.
REPEAT	98	120	LRR 1.
REPEAT	121	144	LRR 2.
REPEAT	145	168	LRR 3.
REPEAT	170	192	LRR 4.
REPEAT	193	216	LRR 5.
REPEAT	217	240	LRR 6.
REPEAT	242	264	LRR 7.
REPEAT	265	288	LRR 8.
REPEAT	289	312	LRR 9.
REPEAT	313	336	LRR 10.
REPEAT	337	360	LRR 11.
REPEAT	361	384	LRR 12.
REPEAT	386	408	LRR 13.
REPEAT	409	432	LRR 14.
REPEAT	433	456	LRR 15.
REPEAT	458	480	LRR 16.
REPEAT	482	504	LRR 17.
REPEAT	505	530	LRR 18.
CARBOHYD	64	64	LRR 19.
CARBOHYD	85	85	LRR 20.
CARBOHYD	96	96	N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD	368	368	N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD	515	515	N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQUENCE	605 AA;	66034 MW;	N-LINKED (GLCNAC. . .) (POTENTIAL).
			F6552z23CBEB918F6 CRC64;

FT REPEAT 194 216 LRR 7.  
 FT REPEAT 217 240 LRR 8.  
 FT REPEAT 242 264 LRR 9.  
 FT REPEAT 266 288 LRR 10.  
 FT REPEAT 289 312 LRR 11.  
 FT REPEAT 313 336 LRR 12.  
 FT REPEAT 337 360 LRR 13.  
 FT REPEAT 361 384 LRR 14.  
 FT REPEAT 386 408 LRR 15.  
 FT REPEAT 409 432 LRR 16.  
 FT REPEAT 433 456 LRR 17.  
 FT REPEAT 458 478 LRR 18.  
 FT REPEAT 479 504 LRR 19.  
 FT REPEAT 506 529 LRR 20.  
 FT REPEAT 543 566 LRR 21.  
 FT CARBOHD 64 64 N'-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHD 85 85 N'-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHD 96 96 N'-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHD 368 368 N'-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHD 515 515 N'-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHD 578 578 N'-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHD 586 586 N'-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 603 AA: 66811 MW: DCD7637D9A5037C CRC64;

Query Match 9.2%; Score 344; DB 1; Length 603;  
 Best Local Similarity 23.3%; PRED. No. 2.6e-17;  
 Matches 141; Conservative 63; Mismatches 194; Indels 208; Gaps 16;  
 QY 8 LLLAWATA-----GATATPVVPPWPHPCPPCOPCACQTRPWYIPRSSYREATVDCN 56  
 Db 13 LLAFAFWWALGPGCHQGLDGPQGASADE---GQCPVACTCS---HDDYTDLSVFC 61  
 QY 57 DLFLTAVPPALPAGTQMLLQNSNTV-----DSEGLY 91  
 Db 62 SKNLTILPDPDIPVSTRALWLDGNNLSSIPSAPQNISSLDFLNLOGSWLRSLEPA 120  
 QY 92 LANLTFLDLSNSFSQARDCCDFHALQQLQSLHLENQNTLRL----- 133  
 Db 121 LQNLNYLHLERNLRLNLAVGLFLHTPSLASLSSNLIGLLEEGIFOGLSHLWDLNLGWN 180  
 QY 134 -----DHSTAGLASIQLYIYHNHNQLYRIAPRAFFSGLSNLRLHNSNLRAIDSRWFEM 187  
 Db 181 SLVVPDPTVFOGLNLHEVLAGNKLYTQPALFCGGLERLDLSRNALRSVKA 240  
 QY 188 LPNLELTIMIGKVNKA-----LDMN----- 208  
 Db 241 LPRQKXKYLDRNLITAVAPGAFGLGKALRMLDLSRNVRAGLMEFTFGLGLHLVIRLAHN 300  
 QY 209 -----FRPLANLRSVLVLAGMNLRETSQYALGQSQLESLSYDNLQARY----- 252  
 Db 301 AIASRPRTFKDLHFLLELOLQHNRTRQGLERTFFGLGQLEVLTLNDNQTEVRYGA 360  
 QY 253 -----PRALEQPGKLKDNLNNQPLQRYGPGFANMHLKEGLNNM 295  
 Db 361 LFNVAVMNLSGNCNSLSPRLPERVFOGDKLQNLHLSLHESCLGHVRLHTFAGLSGIRRLFLRD 419  
 QY 296 ELYVSDKFLAVNLFLKIDTINNPLSLTHPRPHFLQMETMLNNNALS----- 350  
 Db 420 NSISSEQSLAGLSELLELDLTN-RLTHLPROFQGLGHLETLLSNQLTNSA 478  
 QY 351 -----RQ-----QTVELSPNQEVGLHGNPIR 372  
 Db 479 GPLORAFWLDISHNHLLETQAGLFLSSGLGRVRLSIRNNLSQTSFSPGLERLWLWDANPND 538  
 QY 373 CDPYVIR-----WANTGTRVRFED-----OSTLGAEPDIDQLRPLPVREVPFR 414  
 Db 539 CSCPKLRLDFALQNPQGVVPPREVOTVCEGDDCQPVYTYNNITCAGPANVSGLDLRDVSET 598  
 QY 415 EMDHC 420  
 Db 599 HFV-HC '603

RESULT 5  
 ALS\_MOUSE STANDARD; PRT; 603 AA.  
 ID ALS\_MOUSE  
 AC P70389;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Insulin-like growth factor binding protein complex acid labile chain  
 DE precursor (ALS)  
 DE IGFALS OR ALS OR ALBS.  
 GN MUS musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Boisclair Y.R., Seto D., Hsieh S., Hurst K.R., Ooi G.T.;  
 RX MEDLINE-96413591; PubMed-8816745;  
 RA \*Organization and chromosomal localization of the gene encoding the  
 RT mouse acid labile subunit of the insulin-like growth factor binding  
 RT complex.;"  
 RL Proc. Natl. Acad. Sci. U. S. A. 93:10028-10033(1996).  
 CC -I- FUNCTION: MAY HAVE AN IMPORTANT ROLE IN REGULATING THE ACCESS OF  
 CC CIRCULATING IGFS TO THE TISSUES.  
 CC -I- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 kDa WITH  
 CC IGF-I OR IGF-II AND IGF-BP-3 (BY SIMILARITY).  
 CC -I- SUBCELLULAR LOCATION: Extracellular.  
 CC -I- SIMILARITY: CONTAINS 21 LEUCINE-RICH REPEATS (LRR).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration -  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@ib-sib.ch).  
 CC -----  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR00483; LRR\_Gterm.  
 DR InterPro: IPR00372; LRR\_Nterm.  
 DR InterPro: IPR00592; LRR\_out.  
 DR InterPro: IPR00591; LRR\_Typ.  
 DR Pfam: PF00560; LRR; 19.  
 DR Pfam: PF00463; LRRCT; 1.  
 DR Pfam: PF01462; LRRNT; 1.  
 DR PRINTS: PR00019; LEUICHRPT.  
 DR SMART: SM00370; LRR; 4.  
 DR SMART: SM0082; LRRCT; 1.  
 DR SMART: SM00013; LRRNT; 1.  
 DR SMART: SM00369; LRR; 9.  
 DR KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.  
 FT SIGNAL 1 23 BY SIMILARITY.  
 FT CHAIN 24 603 INSULIN-LIKE GROWTH FACTOR BINDING  
 PROTEIN COMPLEX ACID LABILE CHAIN.  
 FT REPEAT 52 73 LRR 1.  
 FT REPEAT 74 96 LRR 2.  
 FT REPEAT 98 120 LRR 3.  
 FT REPEAT 121 144 LRR 4.  
 FT REPEAT 146 168 LRR 5.  
 FT REPEAT 169 192 LRR 6.  
 FT REPEAT 194 216 LRR 7.  
 FT REPEAT 217 240 LRR 8.  
 FT REPEAT 242 264 LRR 9.  
 FT REPEAT 265 288 LRR 10.  
 FT REPEAT 289 312 LRR 11.  
 FT REPEAT 313 336 LRR 12.  
 FT REPEAT 338 360 LRR 13.  
 FT REPEAT 361 384 LRR 14.  
 FT REPEAT 386 408 LRR 15.

SQ	SEQUENCE	567 AA:	63467 MW:	C48643AA73967A7D CRC64;
Query Match		9.1%;	Score 340.5;	DB 1: length 567;
Best Local Similarity		24.6%;	Pred. No. 4.2e-17;	
Matches	132;	Conservative	70;	Mismatches 21;
				Indels 117; Gaps 15;
QY	17	TATPVW---PHVPCPQACQIR-----		----- PWT 42
QY	8	SAVPLPLRQPF--PCPKTCKCVRDAQCSGGSVAHIAELGLIPTNLTHILLFRMDQGIL 65		
QY	43	PRSSYRATVD--CNDLFLTAVPPA-----		-----LPAG-----TQ 72
Db	66	RNHSGSGMTVLRQMLSDHSIADPOTFNDLVLKUFLTRNKISRLPRATLDKMLLE	125	
QY	73	TULLQNSIVRVTQDSELYGTYLNLTEFLDSNSFSDARDCDPHEALPQQLSHFENDLTL 132		
Db	126	QLFLDHNALRDQDQFLQFQQLNLQELGNQNLSPFLPANLSSREIKLDDLSRNRNTHL 185		
QY	133	EDHSFAGLASQKQELYLNHNQYRIPRAFSGSLNLRLHNSNLRAIDSRAFEMPNLE 192		
Db	186	PKGLGQVKLEKLILYNSQJLTSVDSGGLSNIAGALTEFLRLRNHLISVAGFDRGILNL 245		
QY	193	ILMIGGNKVDATLDDMMRFLRPLANLRLSVLAGMLRLRISDYALLEGQLESLSFYDNLARV 252		
Db	246	SILTSQNLLESUPPALLEFLHVSSVSRSLTFLFENPLELPDVLFGEMAGLRELWNLNGTHL 305		
QY	253	PRRALSGQVPGKFLDANKNP-LQRVPGDFANMLHIKEGLJGLNNMELLSIDKFALVNLPE 311		
Db	306	PAAFARNLQSLQGTLGJGRNPRSLAQLRGVFGCILRVLGHT-NALAEELDDALGLG 364		
QY	312	LTKLQDTINNPRUSFHPRAHPLQMETLMNNALSAHLHQTVESLPNQEVGLHGNPI 371		
Db	365	LROVSLRHN-RLRLAPLRTLRLNLSLSEVQLEHNOQELTCDVFAALPQLTQVLGHNPW 423		
QY	372	RCDC----VIEW-----ANATGIVRVTEPQSTL-----CA2PPD1QRLPV 408		
Db	424	4LDCGLWRFLQWLRHIDILGRDEPQCRGEPRAISLFWELLQSDPWCDDP----- 475		
QY	409	REVFRFEMTDCIPLLISPRSPPSLQ-----VAGGESMVLHCRALAEPPEIYV 457		
Db	476	RSLPLDPPTPTEALEAPVPSWLFNSRQSQTAQVLARGES-----PNRRLW 521		
RESULT	7			
SLIT_DROME				
ID	SLIT_DROME	STANDARD;	PRT;	1480 AA.
AC	P24014			
DT	01-MAR-1992 (Rel. 21, Created)			
DT	01-MAR-1992 (Rel. 21, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Slit protein precursor.			
GN	SLI.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Mecopterida;			
OC	Ephydriidae; Drosophilidae; Drosophila.			
OX	[NCBI_TAXID:7227]			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-91099605; PubMed-2176336;			
RA	Rotheberg J.M., Jacobs J.R., Goodman C.S., Artavanis-Tsakonas S.;			
RT	"Slit: an extracellular protein necessary for development of midline glia and commissural axon pathways contains both EGF and LRR domains."			
RT	Genes Dev. 4:2169-2187(1990).			
-1- FUNCTION: NECESSARY FOR DEVELOPMENT OF MIDLINE GLIA AND COMMISSURAL AXON PATHWAYS SLIT MAY INTERACT WITH EXTRACELLULAR MATRIX MOLECULES.				
-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A SHORT FORM: ARE PRODUCED BY ALTERNATIVE SPLICING.				
-1- TISSUE SPECIFICITY: EXCRETED BY THE MIDLINE GLIA CELLS AND EVENTUALLY DISTRIBUTED ALONG THE AXONS.				
CC	-!- SIMILARITY: CONTAINS 7 EGF-LIKE DOMAINS.			
CC	-!- SIMILARITY: CONTAINS 24 LEUCINE-RICH REPEATS (LRR).			
CC	-!- SIMILARITY: CONTAINS C-TERMINAL CYSTEINE KNOT-LIKE DOMAIN.			
CC	-!- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAIN.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X53959; CAA37910.1; -.			
DR	PIR; A36665; A36665.			
DR	HSSP; P00743; ICF.			
DR	FlyBase; FBan003425; sli.			
DR	IPR001512; Asx_hydroxyl.			
DR	InterPro; IPR00359; Cys_knot.			
DR	InterPro; IPR00561; EGF-like.			
DR	InterPro; IPR00742; EGF_2.			
DR	InterPro; IPR00881; EGF_Ca.			
DR	InterPro; IPR001611; LRR.			
DR	InterPro; IPR00483; LRR_Cterm.			
DR	InterPro; IPR00372; LRR_Nterm.			
DR	InterPro; IPR00592; LRR_out.			
DR	InterPro; IPR00591; LRR_typ.			
DR	InterPro; IPR00791; Laminin_G.			
DR	Pfam; PF00007; Cys_knot; 1.			
DR	Pfam; PF00008; EGF; 7.			
DR	Pfam; PF00054; Laminin_G; 1.			
DR	Pfam; PF00560; LRR; 17.			
DR	Pfam; PF01463; LRRC7; 4.			
DR	Pfam; PF01462; LRRT; 4.			
DR	SMART; SM00041; CT; 1.			
DR	SMART; SM00179; EGF_Ca; 2.			
DR	SMART; SM00019; EGF_Like; 5.			
DR	SMART; SM00370; LRR; 4.			
DR	SMART; SM00082; LRRC7; 4.			
DR	SMART; SM00013; LRRT; 4.			
DR	SMART; SM00369; LRR_TYP; 9.			
DR	SMART; SM00282; LamG; 1.			
DR	PROSITE; PS00010; ASX_HIDROXYL; 3.			
DR	PROSITE; PS01185; CTCK_1; 1.			
DR	PROSITE; PS01225; CTCK_2; 1.			
DR	PROSITE; PS00022; EGF_1; 7.			
DR	PROSITE; PS01186; EGF_2; 5.			
DR	PROSITE; PS01187; EGF_Ca; 2.			
DR	PROSITE; PS55025; LAM_G_DOMAIN; 1.			
KW	Neurogenesis; Glycoprotein; Signal; Alternative splicing; BGF-like domain; Repeat; Leucine-rich repeat.			
FT	SIGNAL_	1	36	
FT	CHAIN	37	1480	SILT PROTEIN.
FT	REPEAT	99	122	
FT	REPEAT	123	146	LRR 1.
FT	REPEAT	148	170	LRR 2.
FT	REPEAT	171	194	LRR 3.
FT	REPEAT	195	218	LRR 4.
FT	REPEAT	220	246	LRR 5.
FT	REPEAT	321	344	LRR 6.
FT	REPEAT	345	368	LRR 7.
FT	REPEAT	369	392	LRR 8.
FT	REPEAT	394	416	LRR 9.
FT	REPEAT	417	440	LRR 10.
FT	REPEAT	447	479	LRR 11.
FT	REPEAT	522	545	LRR 12.
FT	REPEAT	546	569	LRR 13.
FT	REPEAT	570	593	LRR 14.
FT	REPEAT	595	617	LRR 15.
FT	REPEAT	618	641	LRR 16.
FT	REPEAT	643	666	LRR 17.
FT	REPEAT	678	701	LRR 18.
FT	REPEAT	720	743	LRR 19.
FT	REPEAT	745	764	LRR 20.

Query Match: 8.6%; Score: 321; DB 1; Length: 560;  
 Best Local Similarity: 26.6%; Pred. No. 1.1e-15;  
 Matches: 141; Conservative: 57; Mismatches: 227; Indexes: 106; Gaps: 16

CC SUBUNIT: Binds MyD88 via their respective TIR domains (By similarity).  
 CC SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC TISSUE SPECIFICITY: Highly expressed in lung. After intraperitoneal injection of lipopolysaccharide, highly expressed

CC	INJURED VASCULAR SURFACES IN THE ARTERIAL CIRCULATION IS A	QY	159 RATSGLSNLRLHNSNLRAIDSRSWFEMPLPNELIMGGNKVADLLDMRERPLANTRSL 218
CC	CRITICAL INITIATING EVENT IN HEMOSTASIS (BY SIMILARITY).	Db	164 NLFLSSLGKVKLVLDSLRNLTHLPOGLAQKQLEKLYNSR--LMSLSSGLANLGL 220
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein.	QY	219 V---LAGNNLREISDYALEGGLOSLESFSYDNLARYPRRAEQVQPLKELDNLINKPLQ 275
CC	--!- SIMILARITY: CONTAINS 14 LEUCINE-RICH REPEATS (LRR).	Db	221 TEIJKERNHLRSTAAGAEDSLGNLSTLTSGLNLESLLPAFLHVSWLTRLTFLNPLEE 280
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	QY	276 VGPQDFAMMLHKELGNNMELVSIKFLVNLPELKUDITNPKLSFTHPRAHLP 335
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	Db	281 LPEVLFGEMAGLRELWLNG-THRLTAAFFNRLSGLQTLGTRNPPLSALPPGMFHGLT 339
CC	use by non-profit institutions as long as its content is in no way	CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>	CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> )
CC	or send an email to license@sb-sib.ch).	CC	or send an email to license@sb-sib.ch).
CC	EMBL; 269594; CA93440.1; -	QY	336 QMETLMNNALSAHHOVTESLPIQEVGLHGNPIR----- 372
DR	InterPro; IPR001611; LRR.	Db	340 ELAVLAHTNAAELPEDALRGIGRLRQVSRLRNLRALPRLFRNLSSLYTVQLEHNL 399
DR	InterPro; IPR00483; LRR_Cterm.	QY	373 -----CDC-----VIRWATGTRVFTIEPOSTLCABPPDL 403
DR	InterPro; IPR00372; LRR_Out.	Db	400 KILPGDWPAALQFLTRVLIGHNPWLCDQGLMFLQWLRHHLIGRDEPQ--CNGPESR 457
DR	InterPro; IPR00592; LRR_Typ.	QY	404 QRLPV-----REVPFREMDHCPLISRSFPLSQ-----VASGSMVL 443
DR	InterPro; IPR00391; LRR; 1.	Db	458 ASLTFWELLQDOWCPSSRGIPDPDPTENALAKADPDTORPNNSQSWAWVOLVARGES-- 514
DR	PRINTS; PRO01463; LRR; 1.	QY	444 HGRALABPEPEPEIW 457
DR	PRINTS; PRO0019; LEURICHRPT.	Db	515 -----PDNRFYW 521
DR	SMART; SM00082; LRRCT; 1.		
DR	SMART; SM00013; LRRNT; 1.		
DR	SMART; SM00169; LRR_Typ; 10.		
KW	platelet; Transmembrane; Glycoprotein; Blood coagulation;		
KW	repeat; Leucine-rich repeat; Cell adhesion; Signal.		
FT	SIGNAL 1		
FT	CHAIN 16		
FT	REPEAT 17		
FT	REPEAT 567		
FT	REPEAT 522		
FT	REPEAT 543		
FT	REPEAT 544		
FT	REPEAT 567		
FT	REPEAT 96		
FT	REPEAT 97		
FT	REPEAT 120		
FT	REPEAT 122		
FT	REPEAT 144		
FT	REPEAT 145		
FT	REPEAT 168		
FT	REPEAT 169		
FT	REPEAT 192		
FT	REPEAT 215		
FT	REPEAT 217		
FT	REPEAT 240		
FT	REPEAT 241		
FT	REPEAT 266		
FT	REPEAT 288		
FT	REPEAT 289		
FT	REPEAT 312		
FT	REPEAT 337		
FT	REPEAT 338		
FT	REPEAT 361		
FT	REPEAT 362		
FT	REPEAT 387		
FT	REPEAT 409		
FT	REPEAT 409		
FT	REPEAT 414		
FT	REPEAT 415		
FT	REPEAT 416		
FT	REPEAT 417		
FT	REPEAT 418		
FT	REPEAT 419		
FT	REPEAT 420		
FT	REPEAT 421		
FT	REPEAT 422		
FT	REPEAT 423		
FT	REPEAT 424		
FT	REPEAT 425		
FT	REPEAT 426		
FT	REPEAT 427		
FT	REPEAT 428		
FT	REPEAT 429		
FT	REPEAT 430		
FT	REPEAT 431		
FT	REPEAT 432		
FT	REPEAT 433		
FT	REPEAT 434		
FT	REPEAT 435		
FT	REPEAT 436		
FT	REPEAT 437		
FT	REPEAT 438		
FT	REPEAT 439		
FT	REPEAT 440		
FT	REPEAT 441		
FT	REPEAT 442		
FT	REPEAT 443		
FT	REPEAT 444		
FT	REPEAT 445		
FT	REPEAT 446		
FT	REPEAT 447		
FT	REPEAT 448		
FT	REPEAT 449		
FT	REPEAT 450		
FT	REPEAT 451		
FT	REPEAT 452		
FT	REPEAT 453		
FT	REPEAT 454		
FT	REPEAT 455		
FT	REPEAT 456		
FT	REPEAT 457		
FT	REPEAT 458		
FT	REPEAT 459		
FT	REPEAT 460		
FT	REPEAT 461		
FT	REPEAT 462		
FT	REPEAT 463		
FT	REPEAT 464		
FT	REPEAT 465		
FT	REPEAT 466		
FT	REPEAT 467		
FT	REPEAT 468		
FT	REPEAT 469		
FT	REPEAT 470		
FT	REPEAT 471		
FT	REPEAT 472		
FT	REPEAT 473		
FT	REPEAT 474		
FT	REPEAT 475		
FT	REPEAT 476		
FT	REPEAT 477		
FT	REPEAT 478		
FT	REPEAT 479		
FT	REPEAT 480		
FT	REPEAT 481		
FT	REPEAT 482		
FT	REPEAT 483		
FT	REPEAT 484		
FT	REPEAT 485		
FT	REPEAT 486		
FT	REPEAT 487		
FT	REPEAT 488		
FT	REPEAT 489		
FT	REPEAT 490		
FT	REPEAT 491		
FT	REPEAT 492		
FT	REPEAT 493		
FT	REPEAT 494		
FT	REPEAT 495		
FT	REPEAT 496		
FT	REPEAT 497		
FT	REPEAT 498		
FT	REPEAT 499		
FT	REPEAT 500		
FT	REPEAT 501		
FT	REPEAT 502		
FT	REPEAT 503		
FT	REPEAT 504		
FT	REPEAT 505		
FT	REPEAT 506		
FT	REPEAT 507		
FT	REPEAT 508		
FT	REPEAT 509		
FT	REPEAT 510		
FT	REPEAT 511		
FT	REPEAT 512		
FT	REPEAT 513		
FT	REPEAT 514		
FT	REPEAT 515		
FT	REPEAT 516		
FT	REPEAT 517		
FT	REPEAT 518		
FT	REPEAT 519		
FT	REPEAT 520		
FT	REPEAT 521		
FT	REPEAT 522		
FT	REPEAT 523		
FT	REPEAT 524		
FT	REPEAT 525		
FT	REPEAT 526		
FT	REPEAT 527		
FT	REPEAT 528		
FT	REPEAT 529		
FT	REPEAT 530		
FT	REPEAT 531		
FT	REPEAT 532		
FT	REPEAT 533		
FT	REPEAT 534		
FT	REPEAT 535		
FT	REPEAT 536		
FT	REPEAT 537		
FT	REPEAT 538		
FT	REPEAT 539		
FT	REPEAT 540		
FT	REPEAT 541		
FT	REPEAT 542		
FT	REPEAT 543		
FT	REPEAT 544		
FT	REPEAT 545		
FT	REPEAT 546		
FT	REPEAT 547		
FT	REPEAT 548		
FT	REPEAT 549		
FT	REPEAT 550		
FT	REPEAT 551		
FT	REPEAT 552		
FT	REPEAT 553		
FT	REPEAT 554		
FT	REPEAT 555		
FT	REPEAT 556		
FT	REPEAT 557		
FT	REPEAT 558		
FT	REPEAT 559		
FT	REPEAT 560		
FT	REPEAT 561		
FT	REPEAT 562		
FT	REPEAT 563		
FT	REPEAT 564		
FT	REPEAT 565		
FT	REPEAT 566		
FT	REPEAT 567		
FT	REPEAT 568		
FT	REPEAT 569		
FT	REPEAT 570		
FT	REPEAT 571		
FT	REPEAT 572		
FT	REPEAT 573		
FT	REPEAT 574		
FT	REPEAT 575		
FT	REPEAT 576		
FT	REPEAT 577		
FT	REPEAT 578		
FT	REPEAT 579		
FT	REPEAT 580		
FT	REPEAT 581		
FT	REPEAT 582		
FT	REPEAT 583		
FT	REPEAT 584		
FT	REPEAT 585		
FT	REPEAT 586		
FT	REPEAT 587		
FT	REPEAT 588		
FT	REPEAT 589		
FT	REPEAT 590		
FT	REPEAT 591		
FT	REPEAT 592		
FT	REPEAT 593		
FT	REPEAT 594		
FT	REPEAT 595		
FT	REPEAT 596		
FT	REPEAT 597		
FT	REPEAT 598		
FT	REPEAT 599		
FT	REPEAT 600		
FT	REPEAT 601		
FT	REPEAT 602		
FT	REPEAT 603		
FT	REPEAT 604		
FT	REPEAT 605		
FT	REPEAT 606		
FT	REPEAT 607		
FT	REPEAT 608		
FT	REPEAT 609		
FT	REPEAT 610		
FT	REPEAT 611		
FT	REPEAT 612		
FT	REPEAT 613		
FT	REPEAT 614		
FT	REPEAT 615		
FT	REPEAT 616		
FT	REPEAT 617		
FT	REPEAT 618		
FT	REPEAT 619		
FT	REPEAT 620		
FT	REPEAT 621		
FT	REPEAT 622		
FT	REPEAT 623		
FT	REPEAT 624		
FT	REPEAT 625		
FT	REPEAT 626		
FT	REPEAT 627		
FT	REPEAT 628		
FT	REPEAT 629		
FT	REPEAT 630		
FT	REPEAT 631		
FT	REPEAT 632		
FT	REPEAT 633		
FT	REPEAT 634		
FT	REPEAT 635		
FT	REPEAT 636		
FT	REPEAT 637		
FT	REPEAT 638		
FT	REPEAT 639		
FT	REPEAT 640		
FT	REPEAT 641		
FT	REPEAT 642		
FT	REPEAT 643		
FT	REPEAT 644		
FT	REPEAT 645		
FT	REPEAT 646		
FT	REPEAT 647		
FT	REPEAT 648		
FT	REPEAT 649		
FT	REPEAT 650		
FT	REPEAT 651		
FT	REPEAT 652		
FT	REPEAT 653		
FT	REPEAT 654		
FT	REPEAT 655		
FT	REPEAT 656		
FT	REPEAT 657		
FT	REPEAT 658		
FT	REPEAT 659		
FT	REPEAT 660		
FT	REPEAT 661		
FT	REPEAT 662		
FT	REPEAT 663		
FT	REPEAT 664		
FT	REPEAT 665		
FT	REPEAT 666		
FT	REPEAT 667		
FT	REPEAT 668		
FT	REPEAT 669		
FT	REPEAT 670		
FT	REPEAT 671		
FT	REPEAT 672		
FT	REPEAT 673		
FT	REPEAT 674		
FT	REPEAT 675		
FT	REPEAT 676		
FT	REPEAT 677		
FT	REPEAT 678		
FT	REPEAT 679		
FT	REPEAT 680		
FT	REPEAT 681		
FT	REPEAT 682		
FT	REPEAT 683		
FT	REPEAT 684		
FT	REPEAT 685		
FT	REPEAT 686		
FT	REPEAT 687		
FT	REPEAT 688		
FT	REPEAT 689		
FT	REPEAT 690		
FT	REPEAT 691		
FT	REPEAT 692		
FT	REPEAT 693		
FT	REPEAT 694		
FT	REPEAT 695		
FT	REPEAT 696		
FT	REPEAT 697		
FT	REPEAT 698		
FT	REPEAT 699		
FT	REPEAT 700		
FT	REPEAT 701		
FT	REPEAT 702		
FT	REPEAT 703		
FT	REPEAT 704		
FT	REPEAT 705		
FT	REPEAT 706		
FT	REPEAT 707		
FT	REPEAT 708		
FT	REPEAT 709		
FT	REPEAT 710		



Db 330 IPDSFLEHLSLCEFLNSRNLRTFARRLSLPCMLDISHNALETLGARALGSLR 389 FT REPEAT 249 272 LRR 8.  
 QY 276 -----VGPGDFFANMLHLKL----- GLNNNEELY 299 FT REPEAT 273 296 LRR 9.  
 Db 390 TLLQGNALRDLPLPPTFANLNLASLQRINLOGNRVSPCGPDRGPGSCVATSGITSLRS 449 FT REPEAT 321 343 LRR 10.  
 QY 300 SIDK-----FALYNLPELTKLDTNTP----- GLNNNEELY 299 FT REPEAT 344 367 LRR 11.  
 Db 450 LVDNTEELLRAGAFLHPT-LTELDLSSNPGLEVATGAGGLEASLEVLAQGNGLMLQV 508 FT REPEAT 370 391 LRR 12.  
 QY 322 -----RLSFIPIRAFHILP-----QMETLMLNNNALSALHQTVESL-PHLQEVG 365 FT REPEAT 395 418 LRR 13.  
 Db 509 DLPCFLCKRNLNLRN-LSHLPATQAVSLEVLDLURNNFSLLPSAMGLETSLRLR 567 FT REPEAT 419 444 LRR 14.  
 QY 366 LHGNIDIRDCVIRW-----ANATGTRVRFIPESTLC 397 FT REPEAT 446 469 LRR 15.  
 Db 568 LQGNPLSLC-QGNGWLAQHQGRVYDADOLIC 600 FT REPEAT 491 513 LRR 16.  
 FT REPEAT 514 537 LRR 17.  
 FT REPEAT 539 552 LRR 18.  
 FT REPEAT 615 638 LRR 19.  
 FT REPEAT 640 663 LRR 20.  
 FT CARBOHYD 196 196 LRR 21.  
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 454 454 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 488 488 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 648 648 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 667 667 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 782 AA; 88561 MW; 311460B2D4527317 CIG64;

RESULT 15  
 ID CHAO\_TRICA STANDARD PRT; 782 AA.  
 AC P82963;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Chaoptin (Photoreceptor cell-specific membrane protein) (Fragment).  
 OS Tribolium castaneum (Red flour beetle).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;  
 OC Cucujiformia; Tenebrionidae; Tribolium.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GA-1;  
 RA Nie W., Stronach B., Panganiban G., Brown S., Denell R.;  
 RT Molecular characterization of tibial and the 3<sup>rd</sup> end of the  
 tribolium homeotic complex.;  
 RL Submitted (2000-07-01) to the EMBL/GenBank/DBJ databases.  
 CC -1 FUNCTION: REQUIRED FOR PHOTORECEPTOR CELL MORPHOGENESIS. MEDIATES  
 HOMOPHILIC CELLULAR ADHESION (BY SIMILARITY).  
 CC -! SUBCELLULAR LOCATION: EXTRACELLULAR SURFACE OF R-CELL PLASMA  
 MEMBRANE (BY SIMILARITY).  
 CC -! SIMILARITY: BELONGS TO THE CHAOPTIN FAMILY.  
 CC -! SIMILARITY: CONTAINS AT LEAST 21 LEUCINE-RICH REPEATS (LRR).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).  
 CC  
 DR EMBL; AF322227; AAK01654.1; -.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR003592; LRR\_out.  
 DR InterPro; IPR003591; LRR\_TYP.  
 DR Pfam; PF00560; LRR; 19.  
 DR PRINTS; PR00019; LEURCHRPT.  
 DR SMART; SM00310; LRR; 6.  
 DR Glycoprotein; Membrane; Repeat; Leucine-rich repeat; Vision.  
 FT NON\_TER 1 1  
 FT REPEAT 43 66 LRR 1.  
 FT REPEAT 67 90 LRR 2.  
 FT REPEAT 93 116 LRR 3.  
 FT REPEAT 117 140 LRR 4.  
 FT REPEAT 141 164 LRR 5.  
 FT REPEAT 165 188 LRR 6.  
 FT REPEAT 224 247 LRR 7.

Query Match  
 Best Local Similarity 7.5%; Score 279; DB 1; Length 782;  
 Matches 145; Conservative 94; Mismatches 233; Indels 112; Gaps 23;  
 Db 106 SDAROCDFHAPOLISLHLBNOLTRLEDHSFAG -- LASICELVYLNHNGYRIRAFSG 163  
 Db 55 RNPDPNSFHRFLRSKVKHLDQNTIEMIHRGTFQDIDHRLTEVTFSPNSVNRQHTFD 114  
 QY 164 LSNLRLHMLNLLRAIDSNNFEMPFNPLMELMIGNNKVDAILDKNFRPLANLRLVAGM 223  
 Db 115 LIDLOQHLDNDRNRIESLERRAFMPLSKLNRKLNKGKIAATAYETFONLPELELDAYN 174  
 QY 224 NREISDYALEGLQSLSESLSFYDN ----- QLARVPRRALLEQVPGL ----- KFLDLNR 270  
 Db 175 SISSDIDNTDQVGSLG -- MEFVNMSHKNVNLNVWAPSPFQFTGIGGLQNIKVLDLF 232  
 QY 271 NPLQVPGDFANM LHLKELGLNNNEELYSIDKALVNLPELKLDTNTP -- RLSFTH 327  
 Db 233 NNTISVAKQFFRPVDSLKQYLGH - NKLNATKDLFGNMPLHQVLDLSHNSLYELDF -- 289  
 QY 328 PRAFHHLPOMETLMNNNALSALHQTVESLPLNQEVGLHGPNTRC -- DCVIRWANATG 385  
 Db 290 -DTFRNTKKLQLWDTSHNRISEIIPNDLFLRGLNRLTVDFSHNRSLSPDNLFR -- ETG- 344  
 QY 386 RVRFIEPOSTILCAEPPDQLQRPVLPREVFPREMIDHCLPLSPRSFPPSLOQASGSMVLHC 445  
 Db 345 -----LBRDVSHNLLGK ----- LPL ----- TSLSLASQTL ----- 371  
 QY 446 RALAEPEPETYWVTPAGLRUPTAHAGR -- RYRVVPEGTELLERRVTAEEAGLYTCV -- AQ 500  
 Db 372 -----SELDISWSNISSL ----- SHGGQALARFKCISWLDSYNSRLQDAGTFKGIPRLAS 422  
 QY 501 NLVGADTKIVSVVGRALIQQPREGQGLERQVQETHPHILLSWV -- TPPNIVSTNL 557  
 Db 423 LNLCNHSQSLTLEELINGLSF ----- QSLEYTLLHNLNDLNSQVAPLSTPNLSSLIA 474  
 QY 558 WSS -----ASSLR --GOCATALARLPRGTHSNITRLL 588  
 Db 475 FNSLUPTVALEVAGNISLRLYLNLDYNDLNSAVPIVTHSTELRLH 518

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

On protein - protein search, using sw model

Run on:

August 31, 2002, 14:36:13 ; Search time 34.24 Seconds

(without alignments)  
3602.377 Million cell updates/sec

Title: US-09-905-088a-245

Perfect score: 3732

Sequence: 1 MRLLYAPLLAWAGATATV..... RKLPLRSEGEATLPPPLSNS 713

Scoring table: BLOSUM2

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTRIMBL19.\*

- 1: sp\_archea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp Rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	1695	45.4	730	4	Q9P231
2	1687	45.4	716	11	Q9P231
3	1655.5	44.4	705	4	Q9P231
4	1656.5	44.4	708	4	Q9P231
5	1649.5	44.2	707	11	Q9P231
6	1648.5	44.2	707	11	Q9P231
7	1625.5	43.6	718	13	Q9P231
8	875.5	23.5	431	4	Q9P231
9	755.5	20.2	273	4	Q9P231
10	434.5	13.3	614	11	Q9P231
11	429.5	13.2	614	4	Q9P231
12	490	13.1	614	6	Q9P231
13	468.5	12.6	606	4	Q9P231
14	484.5	11.8	1991	11	Q9P231
15	438.5	11.7	640	4	Q9P231
16	432.5	11.6	540	5	Q9P231

17	431	11.5	1094	4	Q9BYB8	09byb8 homo sapien
18	430	11.5	1093	4	Q96JAI	Q96jal homo sapien
19	427.5	11.5	532	5	Q96671	Q96671 drosophila
20	420.5	11.3	628	4	Q9BTNO	Q9bt0 homo sapien
21	407	10.9	733	5	Q24250	Q24250 drosophila
22	407	10.9	737	5	Q9VU51	Q9vu51 drosophila
23	406.5	10.9	719	4	Q96NT16	Q96n16 homo sapien
24	401.5	10.8	737	5	Q965M3	Q965m3 caenorhabdi
25	401.5	10.8	881	5	Q965M2	Q965m2 caenorhabdi
26	401.5	10.8	1447	5	Q16779	Q16779 caenorhabdi
27	399.5	10.7	811	4	Q75139	Q75139 homo sapien
28	399	10.7	78	11	Q61974	Q61974 mus musculus
29	398.5	10.7	789	6	Q9BE71	Q9be71 macaca fasciata
30	395.5	10.6	673	11	Q9CZTS	Q9cz5 mus musculus
31	392.5	10.5	792	4	Q9UDT7	Q9udt7 homo sapien
32	388.5	10.4	832	4	Q9ULH4	Q9ulh4 homo sapien
33	386	10.3	492	11	Q99K16	Q99kt6 mus musculus
34	386	10.3	653	4	Q9HBW1	Q9hbw1 homo sapien
35	385	10.3	788	11	Q9CYK3	Q9cyk3 mus musculus
36	378	10.1	649	4	Q96A85	Q96a85 homo sapien
37	377	10.1	1065	4	Q94894	Q94894 homo sapien
38	374.5	10.0	700	4	Q9P244	Q9p244 homo sapien
39	368	9.9	907	11	Q9Z1P4	Q9z1p4 mus musculus
40	361.5	9.7	542	5	Q9NA46	Q9na46 caenorhabdi
41	361	9.7	809	11	Q9DBY4	Q9dby4 mus musculus
42	361	9.7	1173	5	Q9V7JB	Q9v7jb drosophila
43	360.5	9.7	1531	11	Q88279	Q88279 rattus norvegicus
44	360	9.6	918	11	Q99P18	Q99p18 mus musculus
45	360	9.6	907	4	Q75473	Q75473 homo sapien

#### ALIGNMENTS

RESULT 1  
ID Q9P231  
ID Q9P231; PRELIMINARY; PRT; 730 AA.  
AC Q9P231;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE KIAA1497 PROTEIN (FRAGMENT).  
GN KIAA1497.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20277482; PubMed=10819331;  
RA Nagase T, Kikuno R, Ishikawa K., Hiroswa M., Chara O.;  
RT "Prediction of the coding sequences of unidentified human genes.XVII.The complete sequences of 100 new cDNA clones from brain RNA for large proteins in vitro.",  
RT which code for large proteins in vitro.",  
RL DNA Res. 7:143-150(2000).  
DR EMBL; AD040930; BA96021.1; -.  
DR InterPro; IPR03661; FN\_III.  
DR InterPro; IPR03598; Ig\_C2.  
DR InterPro; IPR03066; Ig\_MHC.  
DR InterPro; IPR01011; LRR.  
DR InterPro; IPR00083; LRR\_Cerm.  
DR InterPro; IPR00072; LRR\_Nterm.  
DR InterPro; IPR00392; LRR\_out.  
DR InterPro; IPR03391; LRR\_TYP.  
DR PFM; PR00047; Ig\_1.  
DR PFM; PR00560; LRR; 9.  
DR PFM; PR01463; LRRCT; 1.  
DR PFM; PR0162; LRRNT; 1.  
DR SMART; SM00060; FN3; 1.  
DR SMART; SM00408; IgC2; 1.  
DR SMART; SM00370; LRR; 4.  
DR SMART; SM0082; LRRCT; 1.  
DR SMART; SM00013; LRRNT; 1.

QY	619	---TSCHRALDRPGITAILAVALL	641	Db	313	ATNNPRUYIHPNAFFRFLPKLSSMLNSNALSYHGTIESPLNLKEISIHSNPIRDCV	372
Db	626	HETSTALAAVMGSMFAVISLASYI	652	QY	377	IRWANAGTRVRIEPOSTLCIAEPPDQIQLPVRPREMTHCLPLTSRSPPSIQA	436
RESULT	3			Db	373	IRWANMKTNRMEPMPSLFCVDPERQONTRQHFRDMCMECLPLIAPEPPSINVE	432
ID	043377	PRELIMINARY;	PRT;	QY	437	SGESMVHCRALAEPEPEBIWWTAGLRLTAPAHAGRRYRVYFEGTEFLRRVYAEAGLYT	496
AC	043377;			Db	433	AGSYVSPHCRAPEQPEIWIPTSGOKLPLTDKFVHSE3TLDINGVTPKEGLYT	492
DT	01-JUN-1998	(TREMBlrel. 06, Created)		QY	497	CYAQNLQGADTKIVSVVYGRALLQPGDDEGGELRQVETRHYHILLSWVPPNTSTNL	556
DT	01-DEC-2001	(TREMBlrel. 19, Last sequence update)		Db	493	CIAATNLVGAQDLSVMIKVDGSPQ--DNNGSINTNIKIRDIGANSVLSWKAASSKILKSSV	549
DE	PROBABLE LEUCINE-RICH REPEAT PROTEIN.			QY	557	TWSSASSIRGQATALARLPRGHSTNITRLQIOWEACQVAFADAHTQACVWART	616
GN	RG118007-1.			Db	610	GLHPDQEKEYEKNNTTLMACIGGLGIGIVCLISL-----SPEMNCDDGHS	657
OS	Homo sapiens (Human).			QY	664	-----PLPAPAWMGWSIPSVRVSAPIVLPN	691
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			Db	617	-----EATSCRAIGDRPHIATIALAVLILAAHLGTCGPRKGVGRR-	663
OX	Mammalia; Buthezia; Primates; Catarhini; Homidae; Homo.			Db	658	YVRNYLQKPTFALGELYPPPLNLWEAGKEKTSLSKVATVIGLPTN	703
RN	[1]			QY	693W5	09H3W5	PRELIMINARY;
RP	SEQUENCE FROM N.A.			AC	09H3W5	09H3W5;	PRT;
RC	TISSUE-FIBROBLAST;			DT	01-MAR-2001	(TREMBlrel. 16, Created)	
RA	Murray J., Langston Y., Clarke C.;			DT	01-MAR-2001	(TREMBlrel. 16, Last sequence update)	
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.			DE	01-DEC-2001	(TREMBlrel. 19, Last annotation update)	
DR	EMBL; AC04142; AAC02752; 1.;			DE	PROTEIN-3).		
DR	InterPro; IPR003591; FN3; 1.			GN	DKFZP761K244 OR NLRR-3.		
DR	Pfam; PF00047; Ig; 1.			OS	Homo sapiens (Human).		
DR	Pfam; PF00560; LRR; 9.			OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
DR	Pfam; PF01463; LRRCT; 1.			OX	Mammalia; Buthezia; Primates; Catarhini; Homidae; Homo.		
DR	Pfam; PF01462; LRRN; 1.			RN	[1]	SEQUENCE FROM N.A.	
DR	PRINTS; PR00019; LEURICHRPT.			RC	TISSUE-AMGGDA;		
DR	SMART; SM00408; IgC2; 1.			RA	Bloecher H., Boecker M., Brandt P., Mewes H.W., Weil B., Wiemann S.;		
DR	SMART; SM00370; LRR; 4.			RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.		
DR	SMART; SM00082; LRRCT; 1.			RN	[2]	SEQUENCE FROM N.A.	
DR	SMART; SM00013; LRRN; 1.			RA	Hamano S., Inuzuka H., Morohashi A., Ohira M., Nakagawara A.;		
DR	SMART; SM00369; LRR; 1.			RT	"Human neuronal leucine-rich repeat protein-3(NLRR-3).";		
DR	KW	immunoglobulin domain.		RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.		
DR	SEQUENCE	705 AA;	9049 MW;	Db	EMBL; AL442092; CAC09450; 1.;		
DR	SEQUENCE	705 AA;	9049 MW;	Db	EMBL; AB00967; BAB47184; 1.;		
DR	SEQUENCE	333;	44,48;	Db	EMBL; AB003961; IPR003591; FN3; 1.		
DR	Best Local Similarity	47.2%	; Score 1656.5;	Db	InterPro; IPR003592; LRR_out.		
DR	Matches	333;	Conservative 109;	Db	InterPro; IPR003598; Ig; 1.		
DR	Matches	217;	Indels 47;	Db	InterPro; IPR003006; Ig_MHC.		
QY	17	TATWPPVPPWHPVCPQCACQIQRPMWTPRSYREATVDCNDLFTAVPLPACTQTLI	76	Db	InterPro; IPR001611; LRR.		
QY	14	TTLVQAVDKVDCPRLCTCEIRPWPFTPRYMEASTVDENDGLITFPARLPAWNL	73	Db	InterPro; IPR00483; LRR_Cterm.		
QY	77	OSNSIVRVDQSELGYLANLTEDLSQNSFSDARCCDFHALPOLISLHLEENQLTRLEDHS	136	DR	InterPro; IPR003599; Ig.		
QY	74	OTNNIAKIETS-TDFPVNLTGDLDSQNLNSVNTNPKMOPOLSSVYLENKLTPEKC	132	DR	InterPro; IPR003598; Ig_C2.		
QY	137	FAGLSLQELYLHNQLYRIAPRAFGSLSNLRLHNSNLRAITDSRWEMLPNELEMT	196	DR	InterPro; IPR003591; FN3; 1.		
Db	133	LSELNLQELYLHNLLSPISGAFGLHNLRLHNSNLRAITDSRWEMLPNELEMT	192	DR	InterPro; IPR003592; LRR_Nterm.		
Db	197	GGKRVDAIDMNFRPLANLRLVLAGMILRETSYDALGSGLOSSLSPSYDNLQARVPRRA	256	DR	InterPro; IPR003592; LRR_out.		
Db	193	GEMPIRKDMNPKPLNERSLYTAGINTEPDNALVGENLESISPYDNRILKVPVRA	252	DR	InterPro; IPR003591; FN3; 1.		
QY	257	LEQVPGLKFPLDKNPKLPLQRYGCPDFFANMLKELGNNMELYSIDKFALVNPLPEKID	316	DR	InterPro; IPR003591; LRR_out.		
Db	253	LOKVNVLKFPLDNKPKINRQDFSNMLKELGNNMELYSIDSLAVDNLPELDRKIE	312	DR	PRINTS; PR00019; LEURICHRPT.		
QY	317	ITNNPRLSFHPRPHFQPMETLMNNNNSALHQQTVESLPNLOEVGLHGNPPIRCDCV	376	DR	SMART; SM00409; Ig; 1.		
DR				DR	SMART; SM00408; IgC2; 1.		
DR				DR	SMART; SM00370; LRR; 4.		
DR				DR	SMART; SM00082; LRRCT; 1.		

RESULT 6

ID P97860 PRELIMINARY; PRT; 707 AA.

AC P97860;

DT 01-MAY-1997 (TREMBLrel. 03, Created)

DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)

DE LEUCINE-RICH REPEAT PROTEIN PRECURSOR (FRAGMENT).

GN LRRN3.

OS Mus musculus (Mouse).

RA Taniguchi H., Tohyama M., Takagi T.;

RT "Cloning and expression of a novel gene for a protein with leucine-rich repeats in the developing mouse nervous system.",

RL Brain Res. Mol. Brain Res. 36:45-52(1996).

DR EMBL: D49802; BAA05622.1; -.

DR MGI: 1106036; LRRN3.

DR InterPro: IPR003961; FN\_III.

DR InterPro: IPR003598; Ig\_C2.

DR InterPro: IPR003006; Ig\_MHC.

DR InterPro: IPR01611; LRR.

DR InterPro: IPR00483; LRR\_Cterm.

DR InterPro: IPR00372; LRR\_Out.

DR InterPro: IPR003591; LRR\_typ.

DR Pfam: PF00041; fn3; 1.

DR Pfam: PF00047; fn3; 1.

DR Pfam: PF00560; LRR; 10.

DR Pfam: PF01463; LRRCT; 1.

DR PRINTS; PR00019; LEURICHRPT.

DR SMART; SM00050; FN3; 1.

DR SMART; SM00408; IgC2; 1.

DR SMART; SM00370; LRR; 4.

DR SMART; SM00052; LRRCT; 1.

DR SMART; SM00013; LRRN; 1.

DR SMART; SM00369; LRR\_Typ; 1.

KW Immunoglobulin domain; Signal.

FT SIGNAL\_1 1 20

FT CHAIN 21 >707 POTENTIAL.

FT NON\_TER 707 707 MW; F4E7C1573DD165B CRC64;

SEQUENCE 707 AA; 79156 MW;

RESULT 7

ID 073675 PRELIMINARY; PRT; 718 AA.

AC 073675;

DT 01-AUG-1998 (TREMBLrel. 07, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DE NEURONAL LEUCINE-RICH REPEAT PROTEIN.

GN XNLR-1.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI\_TaxID=8355;

RN 1

RP SEQUENCE FROM N\_A.

RR MEDLINE:99070063; PubMed=9852961;

RR Hayata T., Uochi T., Asashima M.;

RT "Molecular cloning of XNLR-1, a Xenopus homolog of mouse neuronal leucine-rich repeat protein expressed in the developing Xenopus nervous system.",

Gene 221:159-166(1998).

DR EMBL: LAB01462; BAA28530; 1; -.

DR InterPro: IPR003961; FN\_III.

DR InterPro: IPR00398; Ig\_C2.

DR InterPro: IPR003006; Ig\_MHC.

DR InterPro: IPR01611; LRR.

DR InterPro: IPR00483; LRR\_Cterm.

Query Match 44.2%; Score 1648.5; DB 11; Length 707;

Best Local Similarity 47.7%; Pred. No. 2,5e-117;

Matches 328; Conservative 111; Mismatches 225; Indels 23; Gaps 5;

QY 17 TATVPPVPHWVPPQACQIRDWYTPRSYRATTVCNDLFLTAVPAPGQTILL 76

Db 17 TTVQVADKKVDCPQLCCEIRWPTPSIYMASTVDCNDIGLNPAPLDTQILL 76

DT 01-OCT-2000 (TREMBLEL, 15, Created)  
 DT 01-OCT-2000 (TREMBLEL, 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLEL, 19, Last annotation update)  
 DE PLACE106239 PROTEIN (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OS NCBI\_TAXID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RC RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Ishihashi T., Fujimori K.,  
 RA Tanai H., Kimate M., Watabane M., Hiracka S., Ishii S., Kawai Y.,  
 RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,  
 RA Masuho Y., Kaneko K.;  
 RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AK001991; BAA94025.1; -.  
 DR Interpro; IPR001611; LRR.  
 DR Interpro; IPR000372; LRR\_Nterm.  
 DR Interpro; IPR003592; LRR\_out.  
 DR Interpro; IPR003591; LRR\_typ.  
 DR Pfam; PF00560; LRR; 6.  
 DR Pfam; PF0462; LRRN; 1.  
 DR SMART; SM00370; LRR; 2.  
 DR SMART; SM00013; LRRN; 1.  
 DR SMART; SM00369; LRR\_TYP; 1.  
 FT NON\_TER 273 273  
 SQ SEQUENCE 273 AA; 30832 MW; B59F070A6D7D11FF CRC64;

Query Match 20.2%; Score 755.5; DB 4; Length 273;  
 Best Local Similarity 58.8%; Pred. No. 1e-49;  
 Matches 151; Conservative 39; Mismatches 66; Indels 1; Gaps 1;

QY 17 TATVPPVPPWHVPCPPQACOIRPWTPTSSYREATVDCNDLFLTAVPPALPAGTQLL 76  
 DB 17 TTLVQAVDKKVDCPRLCTCIRPWTPTSSYREATVDCSDGLTTPAHLPAWQILL 76

QY 77 QSN\$VVRDQSELGLYANLTTELDI\$QNSFSDARDODFHALPOLISLHLEENQNLTHEDHS 136  
 DB 77 OTNNIAKIEVS-TDPVNPVTLGLDLSQNNLSVTNNVKKMPQLISVYLEENKLPEK 135

QY 137 FAGLALSQELYLNHQLYRATPAPRSGSLNLRHLNSNLRADSWEMLPNLNEITMI 196  
 DB 136 LSELSNLQELYINHNHLISI\$PGARIGHLNLRLHLNSNRLQMINSKWEDALPNEITMI 195

QY 197 GGKVKDAILDMNFRPLANLRSVLAGMNLREISDYALEGQSLSSLSFVNDQLARVPRRA 256  
 DB 196 GENPIIRKDMNFKLPLINRLSIVLGINLTIEPDNALVLENLSSISFYDNRLKPVRA 255

QY 257 LEQFGKLTLNKNPL 273  
 DB 256 LQKVUNLKFLDLNKNP 272

RESULT 10

QD010 ID QD010 PRELIMINARY; PRT; 614 AA.

AC QD010;  
 DT 01-JUN-2001 (TREMBLEL, 17, Created)  
 DT 01-JUN-2001 (TREMBLEL, 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBLEL, 19, Last annotation update)  
 DE ADULT MALE TESTIS cDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,  
 DE CLONE:4930471K13, FULL INSERT SEQUENCE.  
 OS Mus musculus (Mouse);  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.  
 OS NCBI\_TAXID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.

Query Match 13.3%; Score 494.5; DB 11; Length 614;  
 Best Local Similarity 25.6%; Pred. No. 2.9e-29;  
 Matches 173; Conservative 97; Mismatches 251; Indels 155; Gaps 18;

QY 1 MRLLYAPLIAW-----VAGATATVPPVPHVPCPPQACQIRPWTPTSSYREA 50  
 DB 6 MRSMPSPPLACQWPILLVLGVSAT-----GCPRBCS-----AOD 46

QY 51 TVDCNDLFLTAVPPALPAGTQLLQNSIVRVDQSELGLYANLTTELDI\$QNSFSDARD 110  
 DB 47 RAVLCHKRKREVAPEGIPPTETRLDQGKRNKTRINQDDEAFSPHEEHLINENIVSAEP 106

QY 111 CDFH-----ALPQLSLHLEENQNLTHEDISFAGL8OEL 146  
 DB 107 GAFNPLNLNRLTGLRSNRKLIPLGVFTGLSNLTKIDISENKIVTLLDYMFDQLYNKL 166

QY 147 YLNHQLYRATPAPRSGSLNLRHLNSNLRADSWEMLPNLNEITMI 206

DB 167 EVGNDLVLVYTHRAFSGLNSLEQTLTKECNTSITPTEALSHLGLIVRLRHNIAIRD 226

QY 207 MNFPLANLRSVLAGMNLREISDYALEGQSLSSLSFVNDQLARVPRRA 265  
 DB 227 YSFKRLYRKYLESHWPLTWMPLWPNCLYGL-NUTLSFTHCNITAVPLVAVRLVYLF 285

QY 266 LDLNKNPLQVPGDFANMLKELGLNNMELVNSIDKAVLNUPELKLDTINPRSF 325  
 DB 286 LNLSYNPIGIE---GSMHL-----ELLRQELQV-----GGQAV 319

```

1 MRLIVAPLILAW-----VAGATTPVPPWHPCCPACQIRRWWTPRSSRE 50
2 VRSMPSPLACWQPOILLVILGVSLGSAT-----GCPPRCES-----AQD 46
3 51 TTVDCCNDLFLTAVPPALPGTQTLQNSITRVQDQSBGLYVANLTEDLSNSFSDARD 110
4 Db 47 RAVLCHKRRFREVAPPEGPTETRUDLGKRNRIKILNQEFASPFHLEELNENIVSVEP 106
5 QY 111 CDFH-----ALPOLSLHLEENOLTRLDHSFAGLASQEL 146
6 Db 107 GAFPNLNLRTLGLRSNRLKLPLGVFTGLSNLTKLDSNKIVLTDYMFODLYNLKS 166
7 QY 147 YINHNGQLYIARAFSGLSILRHLHNSNLRIDSRWEMPNLEMLIGGNKVIDL 206
8 Db 167 EVGDNDLIVYIHSRAFSGINSLSBOLTLKCNLISPTEAHLHGLVLRRLINNAIRD 226
9 QY 207 MNFRPLANLRSUVLAGM-NLRETSDYALEGLQGOSLESLSFVYDNDLARVPRRALEQVPLKF 265
10 Db 227 YSFKLXVRLKVEISHPRYDLMTPNQYGL NLTISLISITRCHNLTVAPLAYRHLFR 285
11 QY 266 LDLNKNPLQRVGPGDFANMLHKELGNNMEEVLSIDKFALVNLPELTKDITNNPLSF 325
12 Db 286 LNLSYNPSTIE---GSMRH-----BLLRQEIOLV-----GGOLM 319
13 QY 326 IHPRAFHILPQETMLMANNASALHQTVSLELPNQEVGLGQNPTRDCVIRWANATG 385
14 Db 320 VEPYAERGLNLVYLRVNVSGNQLTLEESVHSFVGUNLTLIDSNPLACDCRLWVRRR 379
15 QY 386 RYRFLIEPOSTLCAEPDPLQRLQIPVREVPFREMDHDCLPLISRSF-----PPSLQ 434
16 Db 380 RUMFNROQPT-CATPERVQGKFRKDFP-----DVLLPNVFTCRARRIRDRKAOQF 429
17 QY 435 VASGESNLVHLARALAPPEPELYWPTAGLRLTPAHAGRRYRVPEGLTLELRVTAEEAGL 494
18 Db 430 VDEGHTWOFVCRADGDPAILWLSPR-KHLSVAKNSNGRLAVFPGTLEVRVYAOVQDNG 488
19 QY 495 YTCVAQNLWGDAD-----KTVSVVGRALQGRDGEQGLELVRQE 535
20 Db 489 YLCIAANAGNDMSMPHLHVRSYSPDMWPHQPNKTFAFIPN---QPG-EGEANSIRATV 542
21 QY 536 THPYHI 541
22 Db 543 PPPFDI 548
23 RESULT 13
24 Q9BZ20 PRELIMINARY; PRT; 606 AA.
25 ID 09BZ20;
26 AC 09BZ20;
27 DT 01-JUN-2001 (T-BMBLrel. 17, created)
28 DT 01-JUN-2001 (T-BMBLrel. 17, Last sequence update)
29 DT 01-DEC-2001 (T-BMBLrel. 19, Last annotation update)
30 DE BA43B23.1 (NEURONAL LEUCINE-RICH REPEAT PROTEIN) (CDNA FLJ31810 FIS, CHAIN).
31 DE CLONE NT_2R12005289, weakly similar to CARBOXYPEPTIDASE N 83 KDA
32 DE CHAIN).
33 OS GN Homo sapiens (Human).
34 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
35 OC Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo.
36 OC NCBI_TAXID=9605;
37 RN [1]
38 RN SEQUENCE FROM N.A.
39 RA Babage A.:
40 RL Submitted (JAN-2001) to the EMBL/GenBank/DDJB databases.
41 RP [2]
42 RP SEQUENCE FROM N.A.
43 RA Ishibashi T., Kaneko K., Yosida M., Watanabe S., Ishida S., Ono Y.,
44 RA Hotuda T., Hirokawa S., Murakawa K., Takiguchi S., Kusano J., Orie Y.,
45 RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
46 RA Yamamoto T., Irie R., Otaki T., Sato H., Wakamatsu A., Ishii S.,
47 RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,

```

RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,  
 RA Suzuki Y., Sugano S., Nagahashi K., Masuho Y., Nagai K., Isogai T.;  
 RT "NEDO" human cDNA sequencing project;  
 RL Submitted (OCR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL253716; CAC22713; 1; -.  
 DR EMBL: AK053712; BAB71167; 1; -.  
 DR InterPro: IPR003590; Ig.  
 DR InterPro: IPR003598; Ig\_c2.  
 DR InterPro: IPR003500; Ig\_Like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR000483; LRR\_Cterm.  
 DR InterPro: IPR000372; LRR\_Nterm.  
 DR InterPro: IPR003591; LRR\_TYP.  
 DR Pfam: PF00047; Ig; 1.  
 DR Pfam: PF00560; LRR; 1.  
 DR Pfam: PF01463; LRRCT; 1.  
 DR Pfam: PF01462; LRRNT; 1.  
 DR PRINS; PR00019; LEURICHRPT.  
 DR SMART; SM00409; Ig; 1.  
 DR SMART; SM00408; IgC2; 1.  
 DR SMART; SM00410; Ig\_Like; 1.  
 DR SMART; SM00082; LRRCT; 1.  
 DR SMART; SM00013; LRRNT; 1.  
 DR SMART; SM00369; LRR\_TYP; 10.  
 KW Immunoglobulin domain.  
 SQ SEQUENCE 606 AA; 68065 MW; CB608E281B066B9D CRC64;

Query Match 11.7%; Score 438.5; DB 4; Length 640;  
 Best Local Similarity 25.5%; Pred. No. 5; 8e-25;  
 Matches 148; Conservative 84; Mismatches 231; Indels 117; Gaps 12;

Qy 2 RLLVAPILLAWVAGATAVPPVPPWHVPCPPCACOTRPWTPRSSYREAT\*VDCNDLF 61  
 Db 20 RALEPPLVVLALQQLVAGLYRAOTCPSCS -----NOFSKVICYRKNLR 68

Qy 62 AVPPALPAGTQTLQSNSTVRDSESELGYVIANLTELDSONSFSARDDDPHAPOLLS 121  
 Db 69 EVPDGISTNTRLLNHNQIQIKVMSKHLRHLTQLSRNHRHTEIGAFNGNLNT 128

Qy 122 LHLEMDNLTRELDHFPAGLASLQEYLHNHQYRTRAPAFGLSLNRLHNSNLRAID 181  
 Db 129 LELEFDNLRTTIPNGAVYLSKLELWLRNNTIESPSYANRIPSURRLDGE----- 181

Qy 182 SRWFEMLPNEILMGGNKKDAILDNFRPLANTSLVLAGWNLRBISDVALEGQLES 241  
 Db 182 -----LKRISYISG-----FEGLSNLKYLNLAMCNLRIPN----- 214

Qy 242 LSFYDYNOLARYPVRRALEQVPGKFLDNLNKPLORYGPGDANMLHKEGLNLNMELVI 301  
 Db 215 -----LTPLIKLDDELDLSGNHLSATRPGSITQGLMLQKLMQISD----- 254

Qy 302 DKFALVNLPELTKLITNNRPLSFTHPRAFHLPOMETLMNNNALSALHQQTVESLPLN 361  
 Db 255 -----IOVFERNAFDNLOSSLVEINLAHNNULTLDLFLPPLHL 293

Qy 362 QEVGHGNPIRCDCVIRWANATGTVRFIRPQSTICA---EPPDQLRPLVREVFREMT 417  
 Db 294 ERIHHNPNWCNCILWLS---WWIKDMAPNTACCACRNTPPNPKGRYIGELDQNYFT 350

Qy 418 DHCLPLISPRSFPPSLQVAGESMVLHCRALAEPEPIVWTPACIRLPAHAGRRYV 477  
 Db 351 CYAPPIVEP--PADLNEVGEAAELKCR--STSTSVSITPNTVMTGAYKTRIAVL 406

Qy 478 PEGTLELRVTAEEGLYTCVAQNMVGADPTKTVSVWGRA-----LLQGR 523  
 Db 407 SDGTLNFTNTVQDGMYTCMVSNVGNTTASATLNVTAAATTTPPSYFSFTVETMEPSQ 466

Qy 524 DEGGQELRLQETHYHILLSWVPPNTSTNLWSSASS 563  
 Db 467 DEARTHDDNNVGPPTP---VVDWET--TNYVTSUTPQSTRS 500

Search completed: August 31, 2002, 14:40:33  
 Job time: 260 sec